

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 23:34:11 : Search time 1733 Seconds  
(without alignments)  
8073.337 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342  
Sequence: 1 GCAGAGCTGAGGGGATGTGG.....CTTCAGTAAATATGTACA 342

Scoring table: IDENTITY NUC  
Gapop 10.0 , Gapext 1.0

Searched: 2888711 seqs, 2045481386 residues

Total number of hits satisfying chosen parameters: 5777422

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : GenBank

1: gb\_ba.\*

2: gb\_hg.\*

3: gb\_in.\*

4: gb\_om.\*

5: gb\_ov.\*

6: gb\_pat.\*

7: gb\_ph.\*

8: gb\_pl.\*

9: gb\_pr.\*

10: gb\_ro.\*

11: gb\_sts.\*

12: gb\_sy.\*

13: gb\_un.\*

14: gb\_vl.\*

15: em\_ba.\*

16: em\_fun.\*

17: em\_hum.\*

18: em\_in.\*

19: em\_mu.\*

20: em\_om.\*

21: em\_or.\*

22: em\_ov.\*

23: em\_pat.\*

24: em\_ph.\*

25: em\_pl.\*

26: em\_ro.\*

27: em\_sts.\*

28: em\_un.\*

29: em\_vl.\*

30: em\_htg\_hum.\*

31: em\_htg\_inv.\*

32: em\_htg\_other.\*

33: em\_htg\_mus.\*

34: em\_htg\_pln.\*

35: em\_htg\_rod.\*

36: em\_htg\_man.\*

37: em\_htg\_vrt.\*

38: em\_sy.\*

39: em\_htgo\_hum.\*

40: em\_htgo\_mus.\*

41: em\_htgo\_other.\*

score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB	ID	Description
1	342	100.0	420	6	AR110201	AR110201 Sequence
2	342	100.0	480	6	AR110202	AR110202 Sequence
3	342	100.0	615	6	AR071301	AR071301 Sequence
4	342	100.0	617	9	HUMINSH	L34838 Homo sapien
5	342	100.0	618	6	A47856	L47856 Sequence 1
6	342	100.0	618	6	A98978	A98978 Sequence 1
7	342	100.0	618	6	AX019983	AX019983 Sequence
8	342	100.0	618	6	BD107918	BD107918 Expressio
9	342	100.0	673	9	BC026254	BC026254 Homo sapi
10	342	100.0	934	6	AR126781	AR126781 Sequence
11	342	100.0	934	6	AR202438	AR202438 Sequence
12	340.4	99.5	615	6	AR071299	AR071299 Sequence
13	340.4	99.5	615	6	AR126780	AR126780 Sequence
14	340.4	99.5	615	6	AR202437	AR202437 Sequence
15	241.2	70.5	482	11	G27877	G27877 human STS S
16	224.4	65.6	108265	9	AL133547	AL133547 Human DNA
17	153	44.7	153	6	AR126784	AR126784 Sequence
18	153	44.7	153	6	AR202441	AR202441 Sequence
19	115.6	33.8	108265	9	AL133547	AL133547 Human DNA
20	99	28.9	123	6	AR126783	AR126783 Sequence
21	99	28.9	123	6	AR202440	AR202440 Sequence
22	90	26.3	93	6	AR126785	AR126785 Sequence
23	90	26.3	93	6	AR202442	AR202442 Sequence
24	88	25.7	801	4	AB000201	AB000201 Equus cab
25	86.4	25.3	428	4	S78800	S78800 relaxin B,C
26	84.8	24.8	549	6	A06852	A06852 S.scrofa mR
27	84.8	24.8	549	6	A17335	A17335 S.scrofa mR
28	84.8	24.8	549	6	I02482	I02482 Sequence 3
29	84.8	24.8	733	4	PIGRELX	K01088 Pig relaxin
30	84.8	24.8	752	6	A06051	A06051 Recombinant
31	84.8	24.8	752	6	A16593	A16593 S.scrofa mR
32	84.8	24.8	753	6	E00122	E00122 mRNA coding
33	79.6	23.3	600	4	AF254739	AF254739 Camelus d
34	78	22.8	543	4	AF233688	AF233688 Felis cat
35	73.4	21.5	567	6	I05447	I05447 Sequence 3
36	71.8	21.0	312	6	I00753	I00753 Sequence 13
37	71.8	21.0	501	9	PTBLX4	Z27245 P.troglodyc
38	71.8	21.0	558	6	A06925	A06925 H.sapiens m
39	71.8	21.0	558	6	A17315	A17315 H.sapiens m
40	71.8	21.0	558	6	E00219	E00219 cDNA encodi
41	71.8	21.0	558	6	I00742	I00742 Sequence 2
42	71.8	21.0	558	6	I00749	I00749 Sequence 9
43	71.8	21.0	558	9	HSREL3	X00948 Human mRNA
44	71	20.8	156592	9	AL135786	AL135786 Human DNA
45	70.2	20.5	558	6	A06919	A06919 H.sapiens m

ALIGNMENTS

RESULT 1  
AR110201  
LOCUS AR110201  
DEFINITION Sequence 1 from patent US 6114307.  
ACCESSION AR110201  
VERSION AR110201.1 GI:13826477  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unclassified.  
REFERENCE 1 (bases 1 to 420)  
AUTHORS Jaspers,S.R., Sprugel,K.H., Ren,H.Ping., Humes,J.M. and Conklin,D.C.  
TITLE Methods for stimulating pancreatic islet cell regeneration  
JOURNAL Patent: US 6114307-A 1 05-SEP-2000;

```
FEATURES             Location/Qualifiers
     source             1..420
BASE COUNT           124 a 100 c 98 g 98 t
ORIGIN
Query Match          100.0%; Score 342; DB 6; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATATGCCCCATG 60
Db 76 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATATGCCCCATG 135
Qy 61 CCTGAGAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 120
Db 136 CCTGAGAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 195
Qy 121 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 196 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 255
Qy 181 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
Db 256 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 315
Qy 241 AATATATATCTTTCCGCAAAAAGAGAGTGGAGCTCACAGATTTCATTCATCTGTTGT 300
Db 316 AATATATATCTTTCCGCAAAAAGAGAGTGGAGCTCACAGATTTCATTCATCTGTTGT 375
Qy 301 GAAGTAATTTGTGACGATGAACTTCAGTTAAATATATGTACA 342
Db 376 GAAGTAATTTGTGACGATGAACTTCAGTTAAATATATGTACA 417

RESULT 2
AR110202
LOCUS             AR110202             480 bp      DNA      linear      PAT 14-FEB-2001
DEFINITION       Sequence 3 from patent US 6114307.
ACCESSION       AR110202
VERSION         AR110202.1 GI:12826478
KEYWORDS
SOURCE          Unknown.
ORGANISM        Unclassified.
REFERENCE       1 (bases 1 to 480)
AUTHORS        Jaspers S.R., Sprugel K.H., Ren H.Ping., Humes J.M. and
                Conklin D.C.
TITLE          Methods for stimulating pancreatic islet cell regeneration
JOURNAL        Patent: US 6114307-A 3 05-SEP-2000;
FEATURES
     source             1..480
BASE COUNT       136 a 114 c 121 g 109 t
ORIGIN
Query Match          100.0%; Score 342; DB 6; Length 480;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATATGCCCCATG 60
Db 136 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATATGCCCCATG 195
Qy 61 CCTGAGAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 120
Db 196 CCTGAGAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 255
Qy 121 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 256 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 315
Qy 181 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
```

```
|||||
Db 316 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 375
Qy 241 AAAATAATATCTTTCCGCAAAAAGAGAGTGGAGCTCACAGATTTCATTCATCTGTTGT 300
Db 376 AAAATAATATCTTTCCGCAAAAAGAGAGTGGAGCTCACAGATTTCATTCATCTGTTGT 435
Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATATGTACA 342
Db 436 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATATGTACA 477

RESULT 3
AR071301
LOCUS             AR071301             615 bp      DNA      linear      PAT 18-FEB-2000
DEFINITION       Sequence 3 from patent US 5910480.
ACCESSION       AR071301
VERSION         AR071301.1 GI:7222189
KEYWORDS
SOURCE          Unknown.
ORGANISM        Unclassified.
REFERENCE       1 (bases 1 to 615)
AUTHORS        Koman A., Chassin D. and Bellet D.
TITLE          Protein called epil/placentin, process for the preparation of this
                protein and pharmaceutical composition containing such, DNA coding
                for said Protein
JOURNAL        Patent: US 5910480-A 3 08-JUN-1999;
FEATURES
     source             1..615
BASE COUNT       186 a 145 c 142 g 142 t
ORIGIN
Query Match          100.0%; Score 342; DB 6; Length 615;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATATGCCCCATG 60
Db 182 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATATGCCCCATG 241
Qy 61 CCTGAGAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 120
Db 242 CCTGAGAGACATTCACCCACCACCCAGAGGGTGGCTGCTGAAATCTGGAGTCCCAA 301
Qy 121 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 302 GAAATGGTGTCAACCTCCCAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361
Qy 181 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
Db 362 ATTCCTAATTTGTACAGAGCTGGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 421
Qy 241 AAAATAATATCTTTCCGCAAAAAGAGAGTGGAGCTCACAGATTTCATTCATCTGTTGT 300
Db 422 AAAATAATATCTTTCCGCAAAAAGAGAGTGGAGCTCACAGATTTCATTCATCTGTTGT 481
Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATATGTACA 342
Db 482 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATATGTACA 523

RESULT 4
HUMINSH
LOCUS             HUMINSH             617 bp      mRNA      linear      PRI 20-SEP-1996
DEFINITION       Homo sapiens early placenta insulin-like peptide EPIL (INSL4) mRNA,
                complete cds.
ACCESSION       L34838
VERSION         L34838.1 GI:1220314
KEYWORDS        insulin family; placentin.
SOURCE          Homo sapiens (human)
ORGANISM        Homo sapiens
```

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 617)  
 Chassin,D., Laurent,A., Janneau,J.L., Berger,R. and Bellet,D.  
 Cloning of a new member of the insulin gene superfamily (INSL4) expressed in human placenta  
 Genomics 29 (2), 465-470 (1995)  
 96115599  
 MEDLINE  
 PUBMED 8566396  
 REFERENCE 2 (bases 1 to 617)  
 Koman,Aetal.  
 AUTHORS Patent application, FR 2721033, 13-JUN-1994  
 JOURNAL Unpublished (1994)  
 TITLE  
 FEATURES Location/Qualifiers  
 1. .617  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /cell\_type="trophoblast"  
 /tissue\_type="placenta"  
 /dev\_stage="6 weeks"  
 /tissue\_lib="subtracted cDNA"  
 1. .617  
 /gene="INSL4"  
 106. .525  
 /note="INSL4"  
 /note="early placenta insulin-like peptide"  
 /codon\_start=1  
 /product="Epil"  
 /protein\_id="AAB08516.1"  
 /db\_xref="GI:1220315"  
 /translation="MASLFRSYLPALWLLSOLLRESLAELRGCGPRFGKLLSYCP  
 MPKFTFTTGGWLLSGRKPKEVSTNNKDGQALGTSFIPNLSPELKKPLSEGP  
 SLKKIILSRKRSRHRFPFCCEVICDDGTSVKLCI"  
 BASE COUNT 189 a 144 c 142 g 142 t  
 ORIGIN  
 Query Match 100.0%; Score 342; DB 9; Length 617;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGTCATATGCCCATG 60  
 Db 181 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGTCATATGCCCATG 240  
 Qy 61 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGGAATCTGGACGTCCCAAA 120  
 Db 241 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGGAATCTGGACGTCCCAAA 300  
 Qy 121 GAAATGTCACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180  
 Db 301 GAAATGTCACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 360  
 Qy 181 ATTCCTAATTTGTACAGAGCTGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240  
 Db 361 ATTCCTAATTTGTACAGAGCTGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 420  
 Qy 241 AAAATAATACTTTCCCGAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCGTGTG 300  
 Db 421 AAAATAATACTTTCCCGAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCGTGTG 480  
 Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342  
 Db 481 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 522  
 RESULT 5  
 A47856  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO9534653.  
 A47856  
 ACCESSION  
 VERSION A47856.1 GI:2301742  
 KEYWORDS

unidentified  
 unidentified  
 unclassified.  
 1 (bases 1 to 618)  
 Koman,A., Chassin,D. and Bellet,D.  
 EPIL/PLACENTINE  
 TITLE Patent: WO 9534653-A 1 21-DEC-1995;  
 JOURNAL ROUSSY INST GUSTAVE (FR)  
 COMMENT Other publication FR 2721033 951215.  
 FEATURES Location/Qualifiers  
 1. .618  
 /organism="unidentified"  
 /mol\_type="genomic DNA"  
 /db\_xref="taxon:32644"  
 107. .526  
 /note="unnamed protein product"  
 /codon\_start=1  
 /protein\_id="CAA03084.1"  
 /db\_xref="GI:2301743"  
 /translation="MASLFRSYLPALWLLSOLLRESLAELRGCGPRFGKLLSYCP  
 MPEKFTFTTGGWLLSGRKPKEVSTNNKDGQALGTSFIPNLSPELKKPLSEGP  
 SLKKIILSRKRSRHRFPFCCEVICDDGTSVKLCI"  
 BASE COUNT 189 a 145 c 142 g 142 t  
 ORIGIN  
 Query Match 100.0%; Score 342; DB 6; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.3e-93;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 Qy 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGTCATATGCCCATG 60  
 Db 182 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACATTCGTGTCATATGCCCATG 241  
 Qy 61 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGCTGGAATCTGGACGTCCCAAA 120  
 Db 242 CCTGAGAAGACATTCACACACCCAGGAGGTGGCTGCTGGAATCTGGACGTCCCAAA 301  
 Qy 121 GAAATGTCACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180  
 Db 302 GAAATGTCACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361  
 Qy 181 ATTCCTAATTTGTACAGAGCTGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240  
 Db 362 ATTCCTAATTTGTACAGAGCTGAAGAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 421  
 Qy 241 AAAATAATACTTTCCCGAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCGTGTG 300  
 Db 422 AAAATAATACTTTCCCGAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCGTGTG 481  
 Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342  
 Db 482 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 523  
 RESULT 6  
 A98978  
 LOCUS  
 DEFINITION Sequence 1 from Patent WO9909172.  
 A98978  
 ACCESSION  
 VERSION A98978.1 GI:6781938  
 KEYWORDS  
 SOURCE unidentified  
 ORGANISM unidentified  
 1 (bases 1 to 618)  
 Mock,P. and Bellet,D.  
 AUTHORS METHOD FOR IDENTIFYING AND LOCATING EXPRESSED EPIL PEPTIDES, CODED  
 TITLE BY THE INSL4 GENE AND THEIR USES  
 JOURNAL Patent: WO 9909172-A 1 25-FEB-1999;  
 FEATURES Location/Qualifiers  
 source 1. .618  
 /organism="unidentified"

```
/mol_type="genomic DNA"
/db_xref="taxon:32644"
107..526
/note="unnamed protein product"
/codon_start=1
/protein_id="CAB70586.1"
/db_xref="GI:6781939"
/translation="MASLFRSYLPALWLLSOLLRESLAAELRGCGPRFGKHLISYCP
MPEKFTTTPGGWLLSGRPMVSTNNKDGALGTTSEFIPNLSPCLKPLSEGP
SLKKIILSRKSRGHRFPFCCEVICDDGTSVKLCI"
BASE COUNT 189 a 145 c 142 g 142 t
ORIGIN
Query Match 100.0%; Score 342; DB 6; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTTGGAAACACTTGTCTCATATTGCCCATG 60
DB 182 GCAGAGCTGAGGGGATGTGTCCTCCGATTTGGAAACACTTGTCTCATATTGCCCATG 241
QY 61 CCTGAGAAGACATTCACCAACCCAGGAGGTGCTGCTGAATCTGACGCTCCCAA 120
DB 242 CCTGAGAAGACATTCACCAACCCAGGAGGTGCTGCTGAATCTGACGCTCCCAA 301
QY 121 GAAATGTTGTCACCTCCAAACAAAGATGGACAGCCTTAGGTACGACATCAGAATTC 180
DB 302 GAAATGTTGTCACCTCCAAACAAAGATGGACAGCCTTAGGTACGACATCAGAATTC 361
QY 181 ATTCTTAATTTGTACCAAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATTTGAAG 240
DB 362 ATTCTTAATTTGTACCAAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATTTGAAG 421
QY 241 AAAATAATCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTGATCCATCTGTGT 300
DB 422 AAAATAATCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTGATCCATCTGTGT 481
QY 301 GAAGTAATTTGTACGAGTGAACCTTCAGTTAAATATTGTACA 342
DB 482 GAAGTAATTTGTACGAGTGAACCTTCAGTTAAATATTGTACA 523
RESULT 8
LOCUS BD107918 618 bp DNA linear PAT 18-SEP-2002
DEFINITION Expression of the INSL4 gene in human embryonic bone tissues and
applications.
ACCESSION BD107918
VERSION BD107918.1 GI:23202736
KEYWORDS JP 2002500888-A/1.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 618)
AUTHORS Laurent,A. and Bourez,D.
TITLE Expression of the INSL4 gene in human embryonic bone tissues and
applications
JOURNAL Patent: JP 2002500888-A 1 15-JAN-2002;
INSTITUT GUSTAVE ROUSSY
COMMENT OS Homo sapiens (human)
PN JP 2002500888-A/1
PD 15-JAN-2002
PF 22-JAN-1999 JP 2000528687
PR 23-JAN-1998 FR 98/00715
PI ANN LAURENT,DOMINIQUE BOUREZ
PC C12N15/09,A61K39/395,A61K39/00,A61K39/395,A61K45/00,A61K48/00,
PC A61P19/00,
PC C07H21/00,C07K14/64,C07K16/26,C12N5/10,C12P21/02,C12Q1/68, PC
GOIN33/53,
PC GOIN33/569,C01N33/68,C12N15/00,A61K37/02,C12N5/00 CC
Expression of the INSL4 gene in human embryonic bone tissues CC
and
CC applications Location/Qualifiers
FH Key (107)..(523).
FT CDS Location/Qualifiers
1..618
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
FEATURES
source
189 a 145 c 142 g 142 t
BASE COUNT 189 a 145 c 142 g 142 t
ORIGIN
Query Match 100.0%; Score 342; DB 6; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCAGAGCTGAGGGGATGTGTCCTCCGATTTGGAAACACTTGTCTCATATTGCCCATG 60
DB 182 GCAGAGCTGAGGGGATGTGTCCTCCGATTTGGAAACACTTGTCTCATATTGCCCATG 241
QY 61 CCTGAGAAGACATTCACCAACCCAGGAGGTGCTGCTGAATCTGAGAGTCCCAA 120
DB 242 CCTGAGAAGACATTCACCAACCCAGGAGGTGCTGCTGAATCTGAGAGTCCCAA 301
QY 121 GAAATGTTGTCACCTCCAAACAAAGATGGACAGCCTTAGGTACGACATCAGAATTC 180
DB 302 GAAATGTTGTCACCTCCAAACAAAGATGGACAGCCTTAGGTACGACATCAGAATTC 361
QY 181 ATTCTTAATTTGTACCAAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATTTGAAG 240
DB 362 ATTCTTAATTTGTACCAAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATTTGAAG 421
QY 241 AAAATAATCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTGATCCATCTGTGT 300
DB 422 AAAATAATCTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTGATCCATCTGTGT 481
QY 301 GAAGTAATTTGTACGAGTGAACCTTCAGTTAAATATTGTACA 342
DB 482 GAAGTAATTTGTACGAGTGAACCTTCAGTTAAATATTGTACA 523
RESULT 7
AX019983 618 bp DNA linear PAT 07-SEP-2000
LOCUS AX019983
DEFINITION Sequence 1 from Patent WO9337780.
ACCESSION AX019983
VERSION AX019983.1 GI:10043812
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1
AUTHORS Laurent,A. and Bellet,D.
TITLE Expression of the insl4 gene in human embryonic bone tissues and
applications
JOURNAL Patent: WO 9337780-A 1 29-JUL-1999;
ROUSSY INST GUSTAVE (FR); LAURENT ANNE (FR); BELLET DOMINIQUE (FR)
FEATURES
source
1..618
Location/Qualifiers
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
CDS
107..526
/note="unnamed protein product"
/codon_start=1
/protein_id="CAC07742.1"
/db_xref="GI:10043813"
/translation="MASLFRSYLPALWLLSOLLRESLAAELRGCGPRFGKHLISYCP
MPEKFTTTPGGWLLSGRPMVSTNNKDGALGTTSEFIPNLSPCLKPLSEGP
SLKKIILSRKSRGHRFPFCCEVICDDGTSVKLCI"
BASE COUNT 189 a 145 c 142 g 142 t
```

```

Query Match      100.0%; Score 342; DB 6; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 60
   |||||||
DB 182 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 241
   |||||||

QY 61 CTTGAGAAGACATTCACCCACCCAGGAGGTGGCTGCTGGAATCTGGACGTCGCCAAA 120
   |||||||
DB 242 CTTGAGAAGACATTCACCCACCCAGGAGGTGGCTGCTGGAATCTGGACGTCGCCAAA 301
   |||||||

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGACAGCTTAGGTACGACATCAGAATTC 180
   |||||||
DB 302 GAAATGGTGTCAACCTCCAAACAAGATGGACAGCTTAGGTACGACATCAGAATTC 361
   |||||||

QY 181 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTTGAAG 240
   |||||||
DB 362 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTTGAAG 421
   |||||||

QY 241 AAAATAATACCTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTCCTGTGT 300
   |||||||
DB 422 AAAATAATACCTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTCCTGTGT 481
   |||||||

QY 301 GAAGTAATTTGTACGATGGAATCTCAGTTAAATTTATGTACA 342
   |||||||
DB 482 GAAGTAATTTGTACGATGGAATCTCAGTTAAATTTATGTACA 523
   |||||||

RESULT 9
BC026254
LOCUS
DEFINITION Homo sapiens, insulin-like 4 (placenta), clone MGC:22499
IMAGE:4770395, mRNA, complete cds.
ACCESSION BC026254
VERSION BC026254.1 GI:20070772
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Strausberg, R.
Direct Submission
Submitted (02-APR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: http://mgc.nci.nih.gov
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: CLONTECH
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (ILNL)
DNA Sequencing by: Sequencing Group at the Stanford Human Genome
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcdpaxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.

Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/ILNL at: http://image.llnl.gov
Series: IRAL Plate: 37 Row: m Column: 5
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 4504710.
Location/Qualifiers
1. .673
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="LocusID:3641"
/db_xref="taxon:9606"
/clone="MGC:22499 IMAGE:4770395"

```

```

/tissue_type="Placenta"
/clone_lib="NIH_MGC_79"
/lab_host="DH10B"
/note="Vector: pDNR-LIB"
123. .542
/codon_start=1
/product="insulin-like 4 (placenta)"
/protein_id="AAH26254.1"
/db_xref="GI:20070773"
/translation="MASLFRSYLPALWLLSLLQRESLAARLRCGPRFKHLISYCP
MPKFTFTTGGWLLLESGRPMYSTNNKDGQALGTTSEFIPNLSPELKPLSEGGP
SLKIIILSRKRSRHRDPFCFEVCDGTSVKLCI"
BASE COUNT 228 a 149 c 150 g 146 t
ORIGIN

Query Match      100.0%; Score 342; DB 9; Length 673;
Best Local Similarity 100.0%; Pred. No. 1.3e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 60
   |||||||
DB 198 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 257
   |||||||

QY 61 CTTGAGAAGACATTCACCCACCCAGGAGGTGGCTGCTGGAATCTGGACGTCGCCAAA 120
   |||||||
DB 258 CTTGAGAAGACATTCACCCACCCAGGAGGTGGCTGCTGGAATCTGGACGTCGCCAAA 317
   |||||||

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGACAGCTTAGGTACGACATCAGAATTC 180
   |||||||
DB 318 GAAATGGTGTCAACCTCCAAACAAGATGGACAGCTTAGGTACGACATCAGAATTC 377
   |||||||

QY 181 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTTGAAG 240
   |||||||
DB 378 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTTGAAG 437
   |||||||

QY 241 AAAATAATACCTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTCCTGTGT 300
   |||||||
DB 438 AAAATAATACCTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTCCTGTGT 497
   |||||||

QY 301 GAAGTAATTTGTACGATGGAATCTCAGTTAAATTTATGTACA 342
   |||||||
DB 498 GAAGTAATTTGTACGATGGAATCTCAGTTAAATTTATGTACA 539
   |||||||

RESULT 10
AR126781
LOCUS
DEFINITION Sequence 4 from patent US 6180364.
ACCESSION AR126781
VERSION AR126781.1 GI:14113374
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 934)
AUTHORS Koman, A., Chassin, D. and Bellet, D.
TITLE Protein called epil/placental, process for the preparation of this
protein and pharmaceutical composition containing such, DNA coding
for said protein
JOURNAL Patent: US 6180364-A 4 30-JAN-2001;
FEATURES
Location/Qualifiers
1. .934
/organism="unknown"
BASE COUNT 311 a 188 c 226 g 209 t
ORIGIN

Query Match      100.0%; Score 342; DB 6; Length 934;
Best Local Similarity 100.0%; Pred. No. 1.4e-93;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 60
   |||||||
DB 501 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAACACTTGTGTCATATTCGCCCATG 560
   |||||||

```

QY 61 CTGAGAAGACATTCACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 120  
|||||  
Db 561 CTGAGAAGACATTCACACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 842  
|||||

RESULT 11  
AR202438  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Query Match 100.0%; Score 342; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 560  
|||||  
QY 61 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 120  
|||||  
Db 561 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 842  
|||||

Query Match 100.0%; Score 342; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 560  
|||||  
QY 61 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 120  
|||||  
Db 561 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 842  
|||||

RESULT 12  
AR071299  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Query Match 99.5%; Score 340.4; DB 6; Length 615;  
Best Local Similarity 99.7%; Pred. No. 4.2e-93;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 60  
|||||  
Db 181 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 240  
|||||  
QY 61 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 120  
|||||  
Db 241 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 300  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 301 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 360  
|||||  
QY 181 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 361 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 420  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 421 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 480  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342  
|||||  
Db 481 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 522  
|||||

RESULT 13  
AR126780  
LOCUS  
DEFINITION  
ACCESSION  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM  
REFERENCE  
AUTHORS  
TITLE  
JOURNAL  
FEATURES  
BASE COUNT  
ORIGIN

Query Match 100.0%; Score 342; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 560  
|||||  
QY 61 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 120  
|||||  
Db 561 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 842  
|||||

Query Match 100.0%; Score 342; DB 6; Length 934;  
Best Local Similarity 100.0%; Pred. No. 1.4e-93;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 60  
|||||  
Db 501 GCAGAGCTGAGGGGATGGTCCCGGATTTGGAAACACCTTGCTGTATATGGCCCCATG 560  
|||||  
QY 61 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 120  
|||||  
Db 561 CTGAGAAGACATTCACACACCCAGGAGGGTGGCTGGTAATCTGGAGCTGCCAAA 620  
|||||  
QY 121 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 180  
|||||  
Db 621 GAAATGGTGTCAACCTCCCAACAAGATGACAAGCCTTAGGTACGACATCAGAATTC 680  
|||||  
QY 181 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 240  
|||||  
Db 681 ATTCTTAATTTGTCAACAGAGCTGAAGAACACCTGTCTGAAGGCGAGCCATCATTTGAAG 740  
|||||  
QY 241 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 300  
|||||  
Db 741 AAAATAATTAATTTCCCGCAAAAAGAGAGTGGAGTCAACAGATTTGATCCATTTCTGTGT 800  
|||||  
QY 301 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342  
|||||  
Db 801 GAAATTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 842  
|||||

Query Match 99.5%; Score 340.4; DB 6; Length 615;  
Best Local Similarity 99.7%; Pred. No. 4.2e-93;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCCCATG 60  
|||||  
DB 181 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCCCATG 240

QY 61 CCTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 120  
|||||  
DB 241 CCTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 300

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCCTTAGGTACGACATCAGAATTC 180  
|||||  
DB 301 GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCCTTAGGTACGACATCAGAATTC 360

QY 181 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGTAAGGGGAGCCATCATTTGAAG 240  
|||||  
DB 361 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGTAAGGGGAGCCATCATTTGAAG 420

QY 241 AAAATAATACCTTCCCGCAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCCTGTGT 300  
|||||  
DB 421 AAAATAATACCTTCCCGCAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCCTGTGT 480

QY 301 GAAGTAATTTGTGACGATGGAATTCAGTTAAATATTATGTACA 342  
|||||  
DB 481 GAAGTAATTTGTGACGATGGAATTCAGTTAAATATTATGTACA 522

RESULT 14  
AR202437 AR202437 615 bp DNA linear PAT 20-APR-2002  
LOCUS Sequence 1 from patent US 6362318.  
DEFINITION AR202437  
ACCESSION AR202437  
VERSION AR202437.1 GI:20256976  
KEYWORDS  
SOURCE Unknown.  
ORGANISM Unknown.  
REFERENCE 1 (bases 1 to 615)  
AUTHORS Koman,A., Chassin,D. and Bellet,D.  
TITLE Protein called EpiT/placentin, process for the preparation of this protein and pharmaceutical composition containing such, DNA coding for said protein  
JOURNAL Patent: US 6362318-A 1 26-MAR-2002;  
FEATURES Location/Qualifiers  
source 1..615  
BASE COUNT 187 a 143 c 143 g 142 t

Query Match 99.5%; Score 340.4; DB 6; Length 615;  
Best Local Similarity 99.7%; Pred. No. 4.2e-93;  
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCCCATG 60  
|||||  
DB 181 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATTGCCCCCATG 240

QY 61 CCTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 120  
|||||  
DB 241 CCTGAGAGACATTCACCCACCACCCAGGAGGTGGCTGCTGAATCTGGAGCTGCCAAA 300

QY 121 GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCCTTAGGTACGACATCAGAATTC 180  
|||||  
DB 301 GAAATGGTGTCAACCTCCAAACAAGATGGACAAGCCCTTAGGTACGACATCAGAATTC 360

QY 181 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGTAAGGGGAGCCATCATTTGAAG 240  
|||||  
DB 361 ATTCTTAATTTGTACAGAGCTGAAGAAACCACTGTCTGTAAGGGGAGCCATCATTTGAAG 420

QY 241 AAAATAATACCTTCCCGCAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCCTGTGT 300  
|||||

Db 421 AAAATAATACCTTCCCGCAAAAGAGAGTGGAGCTCACAGATTTGATCCATTCCTGTGT 480

QY 301 GAAGTAATTTGTGACGATGGAATTCAGTTAAATATTATGTACA 342  
|||||  
DB 481 GAAGTAATTTGTGACGATGGAATTCAGTTAAATATTATGTACA 522

RESULT 15  
G27877/c human STS SHGC-34067, sequence tagged site. STS 29-JUN-1996  
LOCUS 482 bp DNA linear  
DEFINITION G27877  
ACCESSION G27877  
VERSION G27877.1 GI:1396596  
KEYWORDS STS; STS sequence; primer; sequence tagged site.  
SOURCE Homo sapiens (human)  
ORGANISM Homo sapiens  
REFERENCE 1 (bases 1 to 482)  
AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
JOURNAL Myers,R.M.  
COMMENT Unpublished (1996)  
Contact: Richard M. Myers  
Stanford Human Genome Center (SHGC)  
Stanford University School of Medicine  
Department of Genetics, M-344, Stanford, CA 94305, USA  
Tel: 4157259687  
Fax: 4157459689  
Email: myers@shgc.stanford.edu

Primer A: TCAGTGAATTTGTCATTGAGAATAC  
Primer B: AAAGAGAAGTGGAGTGCACAG  
STS size: 150  
PCR Profile:  
Initial incubation: 94 degrees C for 90 seconds  
Denaturation: 94 degrees C for 15 seconds  
Annealing: 62 degrees C for 23 seconds  
Polymerization: 72 degrees C for 30 seconds  
PCR Cycles: 30  
Thermal Cycler: Perkin Elmer 9600  
Protocol:  
Template: 25 ng  
Primer: each 1 uM  
dNTPs: each 200 uM  
Tag Polymerase: 0.05 units/ul  
Total Vol: 10 ul

Buffer:  
MgCl2: 2.5 mM  
KCl: 50 mM  
Tris-HCl: 20 mM  
pH: 8.3

Prepared with primer pairs provided by Sandoz, derived from R75651  
-- Washington University/Merck EST sequence.

FEATURES  
source 1..482  
/organism="Homo sapiens"  
/mol\_type="genomic DNA"  
/db\_xref="taxon:9606"  
/map="9"  
STS 1..150  
primer\_bind 1..25  
primer\_bind complement(130..150)  
BASE COUNT 114 a 99 c 121 g 146 t 2 others  
ORIGIN

Query Match 70.5%; Score 241.2; DB 11; Length 482;  
Best Local Similarity 91.5%; Pred. No. 1e-62;  
Matches 324; Conservative 0; Mismatches 18; Indels 12; Gaps 6;

QY 1 GCAGAGCTGAGGGGATGTGTCGCC---GATTGGAAACACTTGTCTGTC---ATATTGC 54

422	GCAACCTGAGGGATGTGTCGCCCGAATTGGGAACACTGTGTCCTATATTGCC	363
55	CCATGCTGTGAGAAGACATT---CACCACCACCCAGAGGGTGGCTGCTGAAATCTGGA	111
362	CCATGCCCTGAGAAGACATTCACCCACCACCCAGGAGGGTGGCTGCTGAAATCTGGA	303
112	CGT-CCCAAAGAAATGGTGTCAACCT-CCAAACAACAAGATGACAAG-CTTTAGGTACG	168
302	CGTCCCAAAGAAATGGTGTCAACCTCCCAACAACAAGATGACAAGCCCTTAGGTACG	243
169	ACATCAGAAATCATTCCTAAATTTCTCACCAGAGCTGNAGAAACCACTGTCGAAGGGCAG	228
242	ACATCAGAAATCATTCCTAAATTTCTCACCAGAGCTGAAGAAACCACTGTCGAAGGGCAG	183
229	CCATCATTTGAAGAAATAATACTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTTGAT	288
182	CCATCATTTGAAGAAATAATACTTTCCCGCAAAAAGAGAGTGGAGCTCACAGATTTGAT	123
289	CCATCTGTGTGGAAGTAATTTGTGACGATGGGAACCTTCAGTTAAATNTGTACA	342
122	CCATTCTGTGTGGAAGTAATTTGTGACGATGGGAACCTTCAGTTAAATNTGTACA	69

Search completed: October 17, 2003, 02:21:35  
Job time : 1737 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 16, 2003, 23:31:46 ; Search time 206 seconds

(without alignments)  
4481.591 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGGCATGTGG.....CTTCAGTTAAATTATGTACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2552756 seqs, 1349719017 residues

Total number of hits satisfying chosen parameters: 5105512

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N\_Geneseq\_19Jun03.\*  
1: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1980.DAT.\*  
2: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1981.DAT.\*  
3: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1982.DAT.\*  
4: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1983.DAT.\*  
5: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1984.DAT.\*  
6: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1985.DAT.\*  
7: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1986.DAT.\*  
8: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1987.DAT.\*  
9: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1988.DAT.\*  
10: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1989.DAT.\*  
11: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1990.DAT.\*  
12: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1991.DAT.\*  
13: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1992.DAT.\*  
14: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1993.DAT.\*  
15: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1994.DAT.\*  
16: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1995.DAT.\*  
17: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1996.DAT.\*  
18: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1997.DAT.\*  
19: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1998.DAT.\*  
20: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA1999.DAT.\*  
21: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2000.DAT.\*  
22: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001A.DAT.\*  
23: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2001B.DAT.\*  
24: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2002.DAT.\*  
25: /SIDSI/gcgdata/geneseq/geneseqn-emb1/NA2003.DAT.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length DB	ID	Description
1	342	100.0	420	19 AAV44663	Znsl coding seque
2	342	100.0	480	19 AAV44664	N-terminally tagge
3	342	100.0	618	17 AAT10275	Human early placen
4	342	100.0	618	20 AAX83561	Human insulin-like
5	342	100.0	618	20 AAX27490	Human insulin-like
6	342	100.0	649	18 AAT68419	Human relaxin-rela
7	84.8	24.8	752	4 AAN30186	Sequence encoding
8	84.8	24.8	752	4 AAN30196	Sequence encoding

C	9	81.4	23.8	2364	23 AAS69640	DNA encoding novel
C	10	81.4	23.8	2364	23 AAS73158	DNA encoding novel
	11	71.8	21.0	475	23 ABV42732	Human prostate exp
	12	71.8	21.0	489	11 AAQ06595	H2 prorelaxin gene
	13	71.8	21.0	495	11 AAQ06596	H2 prorelaxin gene
	14	71.8	21.0	558	5 AAN40086	Sequence encoding
	15	71.8	21.0	558	10 AAN92483	mRNA sequence of h
	16	71.8	21.0	915	16 AAQ76309	Not I-Bam HI fragm
	17	71.8	21.0	1120	23 ABV22241	Human prostate exp
	18	71.8	21.0	1120	23 ABV22244	Human prostate exp
	19	71.8	21.0	1120	23 ABV22256	Human prostate exp
	20	71.8	21.0	1120	23 ABV22268	Human prostate exp
	21	71.8	21.0	1120	23 ABV22272	Human prostate exp
	22	71.8	21.0	1120	23 ABV22275	Human prostate exp
	23	71.8	21.0	1120	23 ABV22277	Human prostate exp
	24	71.8	21.0	1120	23 ABV22278	Human prostate exp
	25	71.8	21.0	1120	23 ABV22280	Human prostate exp
	26	71.8	21.0	1120	23 ABV22281	Human prostate exp
	27	71.8	21.0	1120	23 ABV22284	Human prostate exp
	28	71.8	21.0	1120	23 ABV22285	Human prostate exp
	29	71.8	21.0	1120	23 ABV22286	Human prostate exp
	30	71.8	21.0	1120	23 ABV22287	Human prostate exp
	31	71.8	21.0	1120	23 ABV22288	Human prostate exp
	32	71.8	21.0	1120	23 ABV22289	Human prostate exp
	33	71.8	21.0	1120	23 ABV22290	Human prostate exp
	34	71.8	21.0	1120	23 ABV22292	Human prostate exp
	35	71.8	21.0	1120	23 ABV22298	Human prostate exp
	36	71.8	21.0	1120	23 ABV22304	Human prostate exp
	37	71.8	21.0	1120	23 ABV22305	Human prostate exp
	38	71.8	21.0	1120	23 ABV22315	Human prostate exp
	39	71.8	21.0	1120	23 ABV22317	Human prostate exp
	40	71.8	21.0	1120	23 ABV22321	Human prostate exp
	41	71.8	21.0	1120	23 ABV22323	Human prostate exp
	42	71.8	21.0	1120	23 ABV22332	Human prostate exp
	43	71.8	21.0	1120	23 ABV22333	Human prostate exp
	44	71.8	21.0	1120	23 ABV22335	Human prostate exp
	45	71.8	21.0	1120	23 ABV22336	Human prostate exp

#### ALIGNMENTS

RESULT 1  
AAV44663  
ID AAV44663 standard; cDNA; 420 bp.  
XX  
AC AAV44663;  
XX  
DT 07-OCT-1998 (first entry)  
XX  
DE Znsl coding sequence.  
XX  
KW Znsl; human; placenta; pancreatic islet cell proliferation;  
XX Insulin secretion; diabetes; therapy; ds.  
XX  
OS Homo sapiens.  
XX  
PN W09827210-A1.  
XX  
PD 25-JUN-1998.  
XX  
PF 16-DEC-1997; 97WO-US33326.  
XX  
PR 16-DEC-1996; 96US-0033003.  
XX  
PR (ZYMO ) ZYMOGENETICS INC.  
XX  
PI Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;  
PI Sprugel KH;  
XX  
XX WPI; 1998-362779/31.  
DR P-PSDB; AAW69168.  
XX

PT New isolated protein, Zinsl - obtained from human placenta, which  
PT increases the proliferation of pancreatic islet cells, used for  
PT treating diabetes

XX Claim 1; Page 60-61; 77pp; English.

XX This sequence encodes the human Zinsl protein of the invention. The Zinsl  
CC protein was isolated from human placenta, and is believed to be a new  
CC version of the mature protein of placentalin, having disulphide bonded A  
CC and B chains. The protein can be used for stimulating the proliferation  
CC of pancreatic islets to increase insulin secretory capacity of mammals.  
CC In particular it can be used for the treatment of diabetes. It can also  
CC be used for stimulating in vitro proliferation of pancreatic islet cells.  
CC It can also be used for production of antibodies and in detection and  
CC diagnosis.

SQ Sequence 420 BP; 124 A; 100 C; 98 G; 98 T; 0 other;

Query Match 100.0%; Score 342; DB 19; Length 420;  
Best Local Similarity 100.0%; Pred. No. 9.2e-101;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 60  
DB 76 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 135  
QY 61 CCTGAGAGACATTCACACACCCAGGAGGTGGCTGGATCTGGAGCTGCCCAA 120  
DB 136 CTGAGAGACATTCACACACCCAGGAGGTGGCTGGATCTGGAGCTGCCCAA 195  
QY 121 GAAATGGTCAACCTCCCAACAAAGATGACAGCTTAGGTACGACATCAGAAATC 180  
DB 196 GAAATGGTCAACCTCCCAACAAAGATGACAGCTTAGGTACGACATCAGAAATC 255  
QY 181 ATTCCTAATTTGTACACAGCTGAGAAACACCTGTCTGAAGGCGACCATCATTTGAAG 240  
DB 256 ATTCCTAATTTGTACACAGCTGAGAAACACCTGTCTGAAGGCGACCATCATTTGAAG 315  
QY 241 AAAATAATCTTTCCCGAAAAAGAGAGTGGACGTACAGATTTGATCCCATTCCTGTTGT 300  
DB 316 AAAATAATCTTTCCCGAAAAAGAGAGTGGACGTACAGATTTGATCCCATTCCTGTTGT 375  
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342  
DB 376 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 417

RESULT 2  
AAV44664  
ID AAV44664 standard; cDNA; 480 BP.

XX AAV44664;  
AC  
XX  
DT 07-OCT-1998 (first entry)  
XX  
DE N-terminally tagged Zinsl coding sequence.

XX Zinsl; human; placenta; placentalin; pancreatic islet cell proliferation;  
KW insulin secretion; diabetes; therapy; ds.  
KW

OS Homo sapiens.

XX W09827210-A1.

PN 25-JUN-1998.

XX 16-DEC-1997; 97WO-US23326.

XX 16-DEC-1996; 96US-0033003.

XX (ZYMO ) ZYMOGENETICS INC.

XX Conklin DC, Hoffmann RC, Humes JM, Jaspers SR, Ren HP;  
PI

PI Sprugel KH;

XX WPI; 1998-362779/31.  
DR P-PSDB; AAW69170.

XX New isolated protein, Zinsl - obtained from human placenta, which  
PT increases the proliferation of pancreatic islet cells, used for  
PT treating diabetes

XX Example 1; Page 62-63; 77pp; English.

XX This sequence encodes a N-terminally tagged version of the human Zinsl  
CC protein of the invention. The Zinsl protein was isolated from human  
CC placenta, and is believed to be a new version of the mature protein of  
CC placentalin, having disulphide bonded A and B chains. The protein can be  
CC used for stimulating the proliferation of pancreatic islets to increase  
CC insulin secretory capacity of mammals. In particular it can be used for  
CC the treatment of diabetes. It can also be used for stimulating in vitro  
CC proliferation of pancreatic islet cells. It can also be used for  
CC production of antibodies and in detection and diagnosis.

SQ Sequence 480 BP; 136 A; 114 C; 121 G; 109 T; 0 other;

Query Match 100.0%; Score 342; DB 19; Length 480;  
Best Local Similarity 100.0%; Pred. No. 9.8e-101;  
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 60  
DB 136 GCAGAGCTGAGGGGATGGTCCCGATTGGAAACACTTGTCTCATATTGCCCATG 195  
QY 61 CCTGAGAGACATTCACACACCCAGGAGGTGGCTGGATCTGGAGCTGCCCAA 120  
DB 196 CCTGAGAGACATTCACACACCCAGGAGGTGGCTGGATCTGGAGCTGCCCAA 255  
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGACAGCTTAGGTACGACATCAGAAATC 180  
DB 256 GAAATGGTGTCAACCTCCCAACAAAGATGACAGCTTAGGTACGACATCAGAAATC 315  
QY 181 ATTCCTAATTTGTACACAGCTGAGAAACACCTGTCTGAAGGCGACCATCATTTGAAG 240  
DB 316 ATTCCTAATTTGTACACAGCTGAGAAACACCTGTCTGAAGGCGACCATCATTTGAAG 375  
QY 241 AAAATAATCTTTCCCGAAAAAGAGAGTGGACGTACAGATTTGATCCCATTCCTGTTGT 300  
DB 376 AAAATAATCTTTCCCGAAAAAGAGAGTGGACGTACAGATTTGATCCCATTCCTGTTGT 435  
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342  
DB 436 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 477

RESULT 3  
AAT10275  
ID AAT10275 standard; cDNA; 618 BP.

XX AAT10275;

DT 30-JUL-1996 (first entry)

XX Human early placental insulin-like protein gene.

XX Insulin growth hormone family; early placental insulin like protein;  
KW placentalin; placental tissue; cytotrophoblast; trimer; probe; primer;  
KW amplification; polymerase chain reaction; tyrosine phosphorylation;  
KW cellular protein; growth factor; human; lactation; promoter; PCR;  
KW regeneration; nerve; muscle; skin; bone tissue; ss.

OS Homo sapiens.

XX Key Location/Qualifiers

FT CDS 107..526

/\*tag= a

FT /product= EPIL/placentin  
 FT sig\_peptide 107..157  
 FT /tag= b  
 FT mat\_peptide 158..523  
 FT /tag= c  
 XX  
 PN W09534653-A1.  
 XX  
 XX 21-DEC-1995.  
 XX  
 XX 12-JUN-1995; 95WO-FR00766.  
 XX  
 XX 13-JUN-1994; 94FR-0007191.  
 XX  
 XX (INSR ) INST ROUSSY GUSTAVE.  
 XX  
 XX Bellet D, Chassin D, Koman A;  
 XX  
 XX WPI; 1996-049682/05.  
 DR P-PSDB: AAR89134.  
 XX  
 XX Early placental insulin-like protein, EPIL/placentin - contains  
 PT growth factor-like activity useful for e.g. promoting lactation or  
 PT for regeneration of nerve, muscle, skin or bone tissue  
 XX  
 PS Claim 1; Page 14-15; 25pp; French.

CC This is the nucleotide sequence encoding a novel member of the insulin  
 CC growth hormone family designated early placental insulin like (EPIL)  
 CC protein or placentin. The gene has been found to be expressed  
 CC exclusively in the early placental tissue with a small amount in full  
 CC term placental tissue. No other tissue expresses this gene. The gene  
 CC was obtained from a cDNA library derived from mRNA isolated from  
 CC cytotrophoblasts derived from first trimester placental tissue using,  
 CC as a probe, a fragment of the gene amplified by primers AAT10276-7.  
 CC Although the specific activity of the protein remains to be elucidated,  
 CC it is thought that the protein will induce tyrosine phosphorylation of  
 CC cellular proteins and may have growth factor-like activities e.g. human  
 CC growth factor type 1 or lactation promoter activities. It may also be  
 CC used to regenerate e.g. nerve, muscle, skin or bone tissue.

XX Sequence 618 BP; 189 A; 145 C; 142 G; 142 T; 0 other;

Query Match 100.0%; Score 342; DB 17; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.le-100;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 CGAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 60  
 DB 182 CGAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 241  
 QY 61 CTTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGTGAATCTGGAGCTCCCAAA 120  
 DB 242 CTTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGTGAATCTGGAGCTCCCAAA 301  
 QY 121 GAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTAGACATCAGAAATTC 180  
 DB 302 GAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTAGACATCAGAAATTC 361  
 QY 181 ATTCCTAATTTGTACACAGAGCTGAAGAACCACTGTCTGAAGGCGCCATCATTAAG 240  
 DB 362 ATTCCTAATTTGTACACAGAGCTGAAGAACCACTGTCTGAAGGCGCCATCATTAAG 421  
 QY 241 AAATAAATACTTCCCGCAAAAGAGAGTGGACGCTACAGATTGATTCATTCTGTGT 300  
 DB 422 AAATAAATACTTCCCGCAAAAGAGAGTGGACGCTACAGATTGATTCATTCTGTGT 481  
 QY 301 GAAGTAATTTGTGAGCATGGAACTTCAGTTAAATATGTACA 342  
 DB 482 GAAGTAATTTGTGAGCATGGAACTTCAGTTAAATATGTACA 523

RESULT 4

AAx83561  
 ID AAX83561 standard; cDNA; 618 BP.  
 XX  
 AC AAX83561;  
 XX  
 DT 21-DEC-1999 (first entry)  
 XX  
 DE Human insulin-like 4 cDNA sequence.  
 XX  
 KW Human; insulin-like 4; INSL4; embryonic; bone tissue; ligament;  
 KW early-placental insulin-like protein; EPIL; antibody; probe; primer;  
 KW diagnosis; pathology; differentiation; proliferation; cartilage;  
 KW ossification; osteoporosis; dysplasia; ss.  
 XX  
 OS Homo sapiens.  
 XX  
 FH Key Location/Qualifiers  
 FT CDS 107..526  
 FT /tag= a  
 FT /product= "INSL4"  
 FT /note= "insulin-like 4 protein"  
 XX  
 PN W09337780-A1.  
 XX  
 XX 29-JUL-1999.  
 XX  
 XX 22-JAN-1999; 99WO-FR00137.  
 XX  
 PR 23-JAN-1998; 98FR-0000715.  
 XX  
 PA (INSR ) INST ROUSSY GUSTAVE.  
 XX  
 XX Laurent A, Bellet D;  
 PI  
 XX WPI; 1999-469135/39.  
 DR P-PSDB: AAY26926.  
 XX  
 XX Expression of the INSL4 gene in human embryonic bone tissue and  
 PT ligaments,  
 XX  
 PS Claim 1; Fig 1; 60pp; French.

CC This sequence represents the human insulin-like 4 gene (INSL4) which  
 CC is expressed in human embryonic bone tissue and ligaments. The 139  
 CC amino acid protein encoded by the gene is designated early-placental  
 CC insulin-like (EPIL) protein. The INSL4 gene encodes 3 different EPIL  
 CC proteins designated EPIL 1, 2 or 3. EPIL 1 is a single chain comprising  
 CC amino acids 18-139, EPIL 2 is a 2 chain protein with chain A comprising  
 CC amino acids 115-139 and chain B comprising amino acids 18-58 and EPIL 3  
 CC is a single chain comprising amino acids 59-114. The nucleic acids and  
 CC protein or antibodies against EPIL 1, 2 or 3, and probes or primers for  
 CC INSL4 are useful for the diagnosis of pathology associated with abnormal  
 CC differentiation and/or proliferation of bone tissue or ligaments or  
 CC development of abnormal cartilage and/or abnormal ossification of forming  
 CC bones, e.g. osteoporosis or dysplasia. Compositions capable of modulation  
 CC differentiation, regeneration and/or proliferation of bone tissue and/or  
 CC ligament cells are also useful in treating bone disease.

XX Sequence 618 BP; 189 A; 145 C; 142 G; 142 T; 0 other;

Query Match 100.0%; Score 342; DB 20; Length 618;  
 Best Local Similarity 100.0%; Pred. No. 1.le-100;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 60  
 DB 182 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACCTTGTCTCATATTGCCCATG 241  
 QY 61 CTTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGTGAATCTGGAGCTCCCAAA 120  
 DB 242 CTTGAGAAGACATTCACCAACCCAGGAGGGTGGCTGGTGAATCTGGAGCTCCCAAA 301  
 QY 121 GAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTAGACATCAGAAATTC 180

```

|||||
302 GAATGGTCAACCTCCACACAAGATGACAAAGCCCTTAGTACGACATCAGAAATTC 361
|||||
181 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
|||||
362 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 421
|||||
241 AAAATATATCTTTCCCGCAAAAGAGAAAGTGGACGTCACAGATTTCATCCATCTGTTGT 300
|||||
422 AAAATATATCTTTCCCGCAAAAGAGAAAGTGGACGTCACAGATTTCATCCATCTGTTGT 481
|||||
301 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGTACA 342
|||||
482 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGTACA 523

RESULT 5
AAX27490
ID AAX27490 standard; DNA; 618 BP.
AC AAX27490;
XX
XX
DI 22-JUN-1999 (first entry)
XX
XX Human insulin-like gene 4.
XX
XX INSUL-4; insulin-like gene; EPIL; early placental insulin-like; antibody;
XX vector; probe; hybridisation; tumour; hypoglycaemia; hyperglycaemia;
XX diabetes; cardiovascular; ss.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 107..526
XX FT /*tag= a
XX FT /gene= "insl-4"
XX FT /product= "EPIL"
XX FT /note= "early placental insulin-like peptide"

XX
XX WO9909172-A1.
XX
XX
XX 25-FEB-1999.
XX
XX 12-AUG-1998; 98WO-FR01799.
XX
XX 03-NOV-1997; 97FR-0013802.
XX
XX 14-AUG-1997; 97FR-0010387.
XX
XX (INSR ) INST ROUSSY GUSTAVE.
XX
XX Bellet D, Bidart JM, Troalen F, Mock P;
XX
XX WPI; 1999-181038/15.
XX
XX P-PSDB; AAW99574.
XX
XX EPIL polypeptides encoded by insulin-like gene 4 - and corresponding
XX nucleic acids, antibodies, probes, primers, etc.
XX
XX Claim 1; Fig 1; 119pp; French.
XX
XX
XX This sequence represent the INSUL-4 (insulin-like gene 4) gene encoding
XX an EpiL (early placental insulin-like) polypeptide. The polypeptide,
XX antibodies to the polypeptide, vectors containing the coding sequence and
XX probes derived from the coding sequence, can be used to treat tumours,
XX preferably angioproliferative tumours, especially Kaposi's sarcoma,
XX tumours of the pancreas, liver, uterus or breast, angiosarcomas,
XX glioblastomas, neuroblastomas, rhabdomyosarcomas or leiomyosarcomas;
XX to promote vascularisation of specific tissues; to treat retinopathy,
XX macular degeneration, psoriasis, endometriosis, rheumatoid arthritis,
XX atherosclerosis or hyperthyroidism; to treat post-angioplasty
XX restenosis; to promote or inhibit embryo implantation; to prevent and/or
XX treat disorders directly or indirectly connected with insulin-like
XX activity; to prevent and/or treat disorders directly or indirectly

```

```

CC connected with a dysfunction in carbohydrate metabolism, especially
CC connected with hypo glycaemia or hyperglycaemia, especially gestational
CC diabetes and diabetic complications, especially cardiovascular
CC complications.
XX
XX Sequence 618 BP; 189 A; 145 C; 142 G; 142 T; 0 other;
XX
Query Match 100.0%; Score 342; DB 20; Length 618;
Best Local Similarity 100.0%; Pred. No. 1.1e-100;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACTTGTGTGCATATTGCCCATG 60
DB 182 GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACTTGTGTGCATATTGCCCATG 241
QY 61 CCTGAGAGACATTCACACACACCCAGGAGGGTGGCTGTGGAATCTGGACGTCCCAA 120
DB 242 CCTGAGAGACATTCACACACACCCAGGAGGGTGGCTGTGGAATCTGGACGTCCCAA 301
QY 121 GAATGGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
DB 302 GAATGGTGTCAACCTCCAAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 361
QY 181 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 240
DB 362 ATTCTTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCATTTGAAG 421
QY 241 AAAATATATCTTTCCCGCAAAAGAGAAAGTGGACGTCACAGATTTCATCTGTTGT 300
DB 422 AAAATATATCTTTCCCGCAAAAGAGAAAGTGGACGTCACAGATTTCATCTGTTGT 481
QY 301 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGTACA 342
DB 482 GAAGTAATTTGTGACGATGAACCTTCAGTTAAATATATGTACA 523

RESULT 6
AAT68419
ID AAT68419 standard; cDNA; 649 BP.
XX
XX AAT68419;
XX
XX 24-JUL-1997 (first entry)
XX
XX Human relaxin-related factor-2 (RRF-2) cDNA.
XX
XX Relaxin-related factor-2; RRF-2; placenta; growth factor; ds.
XX
XX Homo sapiens.
XX
XX Key Location/Qualifiers
XX CDS 104..523
XX FT /*tag= a
XX FT sig_peptide 104...181
XX FT /*tag= b
XX FT mat_peptide 182..520
XX FT /*tag= c
XX
XX WO9716549-A2.
XX
XX
XX 09-MAY-1997.
XX
XX 01-NOV-1996; 96WO-US17342.
XX
XX 21-FEB-1996; 96US-0012016.
XX
XX 03-NOV-1995; 95US-0006221.
XX
XX (REGE-) REGENERON PHARM INC.
XX
XX Davis S;
XX
XX WPI; 1997-272118/24.
XX
XX P-PSDB; AAW17676.

```

FT		/tag= b	
FT		/label= Claim 9	
FT	misc_feature	73..546	
FT		/tag= c	
FT		/label= Claim 5	
FT	misc_feature	169..480	
FT		*tag= d	
FT		/label= Claim 10	
XX	EP86649-A.		
XX	24-AUG-1983.		
XX	11-FEB-1983;	83EP-0300714.	
XX	12-FEB-1982;	82AU-0002695.	
PR	11-FEB-1983;	83AU-0011834.	
XX	(FLOR-) FLOREY INST EXP PHY.		
PA	(FLOR-) FLOREY HOWARD INST.		
XX	Hudson PJ, Haley JD, Niall HD, Shine J;		
PI	WPI; 1983-748587/35.		
DR			
XX			
XX			
PT	Genes and DNA transfer vectors for prolrelaxin expression - useful		
PT	in prodn. of porcine relaxin for veterinary and human use		
XX	Claim 2; Page 1; 50pp; English.		
XX	The inventors claim synthetic porcine preprorelaxin and prorelaxin		
CC	and synthetic A, B and C peptide chains of prolaxin, and a gene-for		
CC	expression of porcine preprorelaxin or prorelaxin, and their sub-		
CC	units (see AAN30186). They also claim a double-stranded DNA fragment		
CC	for the expression of the signal peptide chain of porcine		
CC	preprorelaxin comprising a coding strand and a complementary strand		
CC	corresp. to a defined mRNA sequence (see AAN30187-N30194) which		
CC	corresps. to the most homologous regions between the pig and rat cdna		
CC	sequences. A probe (AAN30195) is also claimed.		
CC	(Updated on 25-MAR-2003 to correct PF field.)		
CC	(Updated on 25-MAR-2003 to correct PA field.)		
XX	Sequence 752 BP; 248 A; 143 C; 151 G; 210 U; 0 other;		
SQ			
	Query Match	24.8%; Score 84.8; DB 4; Length 752;	
	Best Local Similarity	49.5%; Pred. No. 3.9e-17;	
	Matches 107; Conservative 27; Mismatches 82; Indels 0; Gaps 0;		
QY	49 TATTGCCCATGCTCGTGTGAGACATTTCACCACCAACCCAGGAGGGTGGCTGCTGAATCT	108	
Db	: : :       : :       : :       : :       : :       : :       : :		
136	GUGGUCUCGCCGUCUCUGGGAAAGAAGCUGCACGCCGGAAGAGCCUCAGCUGAACAACU	195	
QY	109 GSAGCTCCCAGAAATAATGGTGTCAACTCCCAACAAAGATGGACAGCCTTAGGTACG	168	
Db	: : :       : :       : :       : :       : :       : :       : :		
196	GSACCCCGGCAGAACCAUGCCCAUCUCCACCAAAGAUCCAGAAAUUUAAAGAU	255	
QY	169 ACATCAGAAATTCATTCTTAATTTGTACACAGAGCTCAGAAACCACTGTCTGAAGGGCAG	228	
Db	: : :       : :       : :       : :       : :       : :       : :		
236	AUGUUGGAUUUGUUCUAUUUUGCUAATDUUGCCAGAGGCGUAGGCAACAUGUCUGAGAGGCA	315	
QY	229 CCATCATTTGAGAAATAATATCTTTCCCGCAAAAAG	264	
Db	: :       : :       : :       : :       : :       : :		
316	CCAUCACUGAGAGAGCUCACAAACUACAGUACUAAAG	351	
RESULT 8			
AAN30196	ID		
XX	AAN30196 standard; mRNA; 752 BP.		
XX	AAN30196;		
XX			
DT	25-MAR-2003 (updated)		
DT	25-MAY-1992 (first entry)		

```
XX DE Sequence encoding porcine preprorelaxin.
XX KW Relaxin; hormone; ss.
XX OS Sus scrofa domestica.
XX FH Key Location/Qualifiers
XX FT CDS 1..549
XX FT /*tag= a
XX PN EP86649-A.
XX PD 24-AUG-1983.
XX PF 11-FEB-1983; 83EP-0300714.
XX PR 12-FEB-1982; 82AU-0002695.
XX PR 11-FEB-1983; 83AU-0011834.
XX PA (FLOR-) FLOREY INST EXP PHY.
XX PA (FLOR-) FLOREY HOWARD INST.
XX PI Hudson PJ, Halsey JD, Niall HD, Shine J;
XX WI: 1983-748587/35.
XX P-PSDB: AAP30392.
XX Genes and DNA transfer vectors for prorelaxin expression - useful
XX in prodn. of porcine relaxin for veterinary and human use
XX Disclosure; Fig 5; 50pp; English.
XX CC The inventors claim synthetic porcine preprorelaxin and prorelaxin
XX and synthetic A, B and C peptide chains of prorelaxin, and a gene for
XX expression of porcine preprorelaxin or prorelaxin, and their sub-
XX units (see AAN30186). They also claim a double-stranded DNA fragment
XX for the expression of the signal peptide chain of porcine
XX preprorelaxin comprising a coding strand and a complementary strand
XX corresp. to a defined mRNA sequence (see AAN30187-N30194) which
XX corresponds to the most homologous regions between the pig and rat cDNA
XX sequences. A probe (AAN30195) is also claimed.
XX CC (Updated on 25-MAR-2003 to correct PF field.)
XX CC (Updated on 25-MAR-2003 to correct PA field.)
XX SQ Sequence 752 BP; 247 A; 143 C; 151 G; 211 U; 0 other;
Query Match 24.8%; Score 84.8; DB 4; Length 752;
Best Local Similarity 49.5%; Pred. No. 3.9e-17;
Matches 107; Conservative 27; Mismatches 82; Indels 0; Gaps 0;
QY 49 TATTGCCCCCATGCTGAGAGAGATTCCACCACCACCCAGGAGGGTGGCTGCTGGAATCT 108
DB 136 UGUGGCUCCGUCUCCUGGGAAGNACUCUCUCCAGCCUGAAGCCUCAGCUGGAACU 195
QY 109 GGAGCTCCCAAGAAATGGTGTCACTCCACAAACAAAGATGGACAAGCCTTAGGTACG 168
DB 196 GGACCCCGGCGAAGAACCAUGCCAUCCUCCAUCCAAAGACGAGAAAUUUAAAGAUG 255
QY 169 ACATCAGATTTCCTTAATTGTCACAGAGCTGAAGAAACCATGCTGTGAAGGCAG 228
DB 256 AUGUUGAAGUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 315
QY 229 CCATCATTTGAAGAAATTAATATTCTTCCCGCAAAAAG 264
DB 316 CCAUCACUGAGAGAGCUCACACACAUUCUGCAUAAAG 351
RESULT 9
AAS69640/c
ID AAS69640 standard; cDNA; 2364 BP.
XX AC AAS69640;
```

```
XX DT 13-FEB-2002 (first entry)
XX DE DNA encoding novel human diagnostic protein #5444.
XX KW Human; chromosome mapping; gene mapping; gene therapy; forensic;
XX KW food supplement; medical imaging; diagnostic; genetic disorder; ss.
XX OS Homo sapiens.
XX PN WO200175067-A2.
XX PD 11-OCT-2001.
XX PF 30-MAR-2001; 2001WO-US08631.
XX PR 31-MAR-2000; 2000US-0540217.
XX PR 23-AUG-2000; 2000US-0649167.
XX PA (HYSE-) HYSEQ INC.
XX PI Drmanac RT, Liu C, Tang YT;
XX WI: 2001-639362/73.
XX P-PSDB: ABG05453.
XX New isolated polynucleotide and encoded polypeptides, useful in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits and to assess
XX biodiversity.
XX Claim 1; SEQ ID No 5444; 103pp; English.
XX CC The invention relates to isolated polynucleotide (I) and
XX polypeptide (II) sequences. (I) is useful as hybridisation probes,
XX polymerase chain reaction (PCR) primers, oligomers, and for chromosome
XX and gene mapping, and in recombinant products of (II). The
XX polynucleotides are also used in diagnostics as expressed sequence tags
XX for identifying expressed genes. (I) is useful in gene therapy techniques
XX to restore normal activity of (II) or to treat disease states involving
XX (II). (II) is useful for generating antibodies against it, detecting or
XX quantitating a polypeptide in tissue, as molecular weight markers and as
XX a food supplement. (II) and its binding partners are useful for medical
XX imaging of sites expressing (II). (I) and (II) are useful for treating
XX disorders involving aberrant protein expression or biological activity.
XX The polypeptide and polynucleotide sequences have applications in
XX diagnostics, forensics, gene mapping, identification of mutations
XX responsible for genetic disorders or other traits to assess biodiversity
XX and to produce other types of data and products dependent on DNA and
XX amino acid sequences. AAS64197-AAS94564 represent novel human
XX diagnostic coding sequences of the invention.
XX CC Note: The sequence data for this patent did not appear in the printed
XX specification, but was obtained in electronic format directly from WIPO
XX at ftp.wipo.int/pub/published_pct_sequences.
XX SQ Sequence 2364 BP; 861 A; 537 C; 439 G; 527 T; 0 other;
Query Match 23.8%; Score 81.4; DB 23; Length 2364;
Best Local Similarity 74.1%; Pred. No. 8.1e-16;
Matches 103; Conservative 0; Mismatches 36; Indels 0; Gaps 0;
QY 117 CAAGAATAATGTGTCAACCTCCACACAAAGATGGACAAGCCTTAGGTAGCATCAGA 176
DB 351 CACAGAGATGATGTCATCCACCCTAATAAGATGCAGGACCTTAGATATGACGTGGG 292
QY 177 ATTCATTCCTAATTTGTCACAGAGCTGAAGAAACCACTGTCTGAAGGGAGCATCAT 236
DB 291 ATTCATTCCTAATTTGTCACAGAGCTGAAGAAACCAATTCGTCTGAAGGGTAGCTG 232
QY 237 GAAGAAATAATACTTTCC 255
DB 231 GAGAAAGTACTACTTTCC 213
```

[illegible]

	QY	142	RACANAGATGCAACGAGCCTTAGTGTACGACATCAGAATTTCCTAATTGTGCACCAGAG	201
	DB	162	RACAAGATACAGAAACCATTAATATGATGTGCAATTTGTTGCTAATTTGCCACAGGAG	221
	QY	202	CTGAAGAACCCTGCTGTGAAGGGCAGCCATCATTT	236
	DB	222	CTGAAGTTTAAGCCTGTCTGAGATCGAGCCAGCATTT	256

RESULT 12  
AAQ06595  
ID AAQ06595 standard; DNA: 489 BP.  
XX  
XX AC AAQ06595;  
AC  
DT 27-FEB-1991 (first entry)  
XX  
XX  
DT  
DE H2 prorelaxin gene insert in plasmid pTRIPRORELASP.  
XX

KW Relaxin; PR; ds.  
XX  
XX Homo sapiens.  
OS  
XX WO9013659-A.  
PN  
XX  
XX PD 15-NOV-1990..  
XX  
PF 16-APR-1990; 90WO-US02085.  
XX  
XX 04-MAY-1989; 89DS-0347550.  
PR  
XX  
XX (GETH ) GENENTECH INC.  
PA  
XX Henner DJ, Vandlen RI, Wikins JA, Yansura D;  
PI  
XX WPI; 1990-361494/48.  
XX DR P-SDB; AAR07967.  
XX  
XX  
PT Cleaving polypeptide into polypeptide cleavage prods. - by  
PT treating free-cysteine form of polypeptide at desired junction.  
PS Disclosure; Fig 2A; 56pp; English.

The plasmid was prep'd. from a clone isolated from a cDNA library prep'd. from RNA isolated from human corpus luteum, screened with an H1-cDNA probe. A fragment encoding the N-terminal of PR was isolated and was ligated with a 410 bp fragment encoding AAS 17-153 of PR, and a fragment from pGHG207141 (including the Trp promoter, AAS 1-137 of methGHG, and amp and tet resistance. A portion of the resulting construction (pPRproH2) was ligated with a fragment from the original clone encoding the PR C-terminal, and a fragment from pBR322XAP encoding part of the beta-lactamase gene. This produced PR gene and pPRproRel. A StuI signal sequence was then fused to the PR gene into a vector identical to pTRPStuIRGH in which the HGH gene had been removed. The resulting construction, pTRPStuIPRrel was treated to remove the StuI sequence and the first 11 AAs of H2 PR. This was replaced with a synthetic DNA duplex encoding the first 12 AAs of H2 PR (including Asp1), to produce pPRproRelAsp, the coding part of which is shown below. The plasmid can be used to construct vectors that encode Asp-inserted H2PR with enhanced acid cleavage sites.

See also AAQ06596.

Sequence 489 BP: 156 A; 104 C; 103 G; 126 T; 0 other;  
XX

Query Match 21.0%; Score 71.8; DB 11; Length 489;  
Best Local Similarity 66.5%; Pred. No. 5.3e-13;  
Matches 103; Conservative 0; Mismatches 52; Indels 0; Gaps 0;

QY	82	ACCCAGGAGGGTGGCTGTGGAATCTGGACGTCCCAGAANAATGGTGTAACCTCCAAC	141
		_ _	
Db	103	AGCAGGAAGATGTCCTCGACACACTAGACCAGTGGCAGAAATTGTGCCATCCTTCATC	162

142	QY	AACAAAGATGACCAAGCCCTTAGTACGACATCATCAATTCCTAAATTTGTCCACGAG	201
163	Db	AACAAAGATGACCAAGCCCTTAGTACGACATCATCAATTCCTAAATTTGTCCACGAG	222
202	QY	CTGAAGAAACCACCTGCTCTGAAGGGCAGCCATCATTT	236
223	Db	CTGAAGTTAAACCTGCTCTGAGATGACGCCAGCATTT	257

AAQ06396  
ID AAQ06596 standard; DNA; 495 BP.

AC AA006596:

DT 27-FEB-1991 (first entry)

DE H2 prorelaxin gene insert in plasmid pTR411.

KW Relaxin; PR; ds.

XX Homo sapiens.

XX PN WO9013659-A.

XX PD 15-NOV-1990.

XX  
PF 16-APR-1990:XX  
PR 04-MAY-1989: 89US-0347550-

XX PA (GETH ) GENENTECH INC

XX  
PT  
Henner D.T. vanden R.

XX  
DB  
WPI: 1990-361494/48

DR P-PSDB; AAR07988.  
XX

PT	Cleaving polypeptide
PT	treating free-cysteine

XX  
PS  
Disclosure: Fia 2A:

xx The plasmid was prepd. from three plasmids, the parental plasmid,  
cc pPrproreAsp (see AAQ0595 for details), and pPr390 and pPr400-20.  
cc Plasmid pPr390-7 was designed to introduce an Asp codon in the met-  
cc PR gene between the end of the B-chain coding region and the start  
cc of the C-chain encoding DNA. Plasmid pPr400-20 was designed to  
cc introduce an Asp codon in the met-PR gene between the end of the C-  
cc chain (Arg137) and the beginning of the A-chain (Gln138) encoding  
cc regions. Plasmid pPr411 was constructed by ligating together the  
cc regions.  
cc

CC following:  
CC (a) pTRP0relAsp from which the 410 bp BssHII-BglII fragment was  
CC removed (it therefore contained codons for AAs 1-8 and 156-161 of  
CC PR);  
CC (b) a fragment from pTR390-7 contg. codons for AAs19-97, with an  
CC adnl. Asp codon between AAs 33 and 34; and  
CC (c) a fragment from pTR400-20 contg. codons 99-155 of met-PR, with  
CC an extra Asp between codons 137 and 138. The plasmid can be used  
CC to construct vectors that encode Asp-inserted H2PR with enhanced  
CC acid cleavage sites.  
CC See also AAC06595.  
CC

Sequence 495 BP: 158 A; 104 C; 105 G; 128 T; 0 other;  
XX SQ

Query Match	21.0%	Score 71.8;	DB 11;	Length 495;
Best Local Similarity	66.5%	Pred. No. 5.3e-13;		
Matches 103; Conservative	0;	Mismatches 52;	Indels	0

Qy	82	ACCCAGGAGGGTGGCTGCTGGAATCTGGACGTCCCAAGAAATGGTGTCAACCTCCAAC	141
		- -	
Db	106	AGCCAGGAAGATGTCCTCAGACACACTAGACCAGTGGCAGAAATGTGCCATCCTTCATC	165





Search completed: October 17, 2003, 01:52:32  
Job time : 207 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2003, 01:37:56 ; Search time 52 Seconds  
(without alignments)  
2902.943 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGATGTTGG.....CTTCACTTAATATGATACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 569978 seqs, 220691566 residues

Total number of hits satisfying chosen parameters: 1139956

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Lining first 45 summaries

Database : Issued Patents NA.\*  
1: /cgn2.6/ptodata/1/ina/5A.COMB.seq.\*  
2: /cgn2.6/ptodata/1/ina/5B.COMB.seq.\*  
3: /cgn2.6/ptodata/1/ina/6A.COMB.seq.\*  
4: /cgn2.6/ptodata/1/ina/6B.COMB.seq.\*  
5: /cgn2.6/ptodata/1/ina/PCTUS.COMB.seq.\*  
6: /cgn2.6/ptodata/1/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	342	100.0	420	3	US-08-991-890-1
2	342	100.0	480	3	US-08-991-890-3
3	342	100.0	615	2	US-08-482-842B-3
4	342	100.0	934	3	US-09-174-465D-4
5	342	100.0	934	4	US-09-599-564A-4
6	340.4	99.5	615	2	US-08-482-842B-1
7	340.4	99.5	615	3	US-09-174-465D-1
8	340.4	99.5	615	4	US-09-599-564A-1
9	153	44.7	153	3	US-09-174-465D-9
10	153	44.7	153	4	US-09-599-564A-9
11	99	28.9	123	3	US-09-174-465D-7
12	99	28.9	123	4	US-09-599-564A-7
13	90	26.3	93	3	US-09-174-465D-11
14	90	26.3	93	4	US-09-599-564A-11
15	71.8	21.0	489	6	5464756-17
16	71.8	21.0	495	6	5464756-19
17	71.8	21.0	915	1	US-08-443-568B-11
18	71.8	21.0	915	5	PCT-US94-06997-11
C 19	69	20.2	231	4	US-09-439-313-465
C 20	69	20.2	231	4	US-09-352-616A-465
C 21	38.6	11.3	2574	3	US-09-142-529-2
C 22	38.6	11.3	2574	4	US-10-045-428A-2
C 23	35	10.2	116592	4	US-09-818-512-3
C 24	34.4	10.1	1001	4	US-09-641-638-123
C 25	31	9.1	2585	3	US-09-008-697A-7
C 26	30.2	8.8	2895	4	US-09-422-936-52
C 27	30	8.8	932	4	US-09-257-179-20

C 28	30	8.8	2250	1	US-08-483-639-1	Sequence 1, Appl
29	30	8.8	5093	1	US-08-468-036-23	Sequence 23, Appl
30	30	8.8	5093	2	US-08-376-843-23	Sequence 23, Appl
C 31	29.6	8.7	9069	4	US-08-961-527-37	Sequence 37, Appl
32	29.4	8.6	553	2	US-08-721-488-7	Sequence 7, Appl
33	29.4	8.6	1789	4	US-08-936-165A-116	Sequence 116, App
C 34	29.4	8.6	1875	2	US-09-511-881A-11	Sequence 11, Appl
C 35	29.4	8.6	2451	2	US-08-820-170A-29	Sequence 29, Appl
C 36	29.4	8.6	2451	3	US-09-055-699-29	Sequence 29, Appl
C 37	29.4	8.6	2451	3	US-09-273-565-29	Sequence 29, Appl
C 38	29.4	8.6	2451	4	US-09-565-538-29	Sequence 29, Appl
C 39	29.4	8.6	2451	4	US-09-661-468-29	Sequence 29, Appl
C 40	29.4	8.6	2451	4	US-09-976-155-29	Sequence 29, Appl
C 41	29.4	8.6	2487	2	US-08-820-170A-32	Sequence 32, Appl
C 42	29.4	8.6	2487	3	US-09-055-699-32	Sequence 32, Appl
C 43	29.4	8.6	2487	3	US-09-273-565-32	Sequence 32, Appl
C 44	29.4	8.6	2487	4	US-09-565-538-32	Sequence 32, Appl
C 45	29.4	8.6	2487	4	US-09-661-468-32	Sequence 32, Appl

## ALIGNMENTS

RESULT 1  
US-08-991-890-1  
; Sequence 1, Application US/08991890  
; Patent No. 6114307  
; GENERAL INFORMATION:  
; APPLICANT: Jaspers, Stephen R.  
; APPLICANT: Sprugel, Katherine H.  
; APPLICANT: Ren, Hong Ping  
; APPLICANT: Humes, Jacqueline M.  
; APPLICANT: Hoffman, Ross C.  
; APPLICANT: Conklin, Darrell C.  
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR  
; NUMBER OF SEQUENCES: 7  
; TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: ZymoGenetics, Inc.  
; STREET: 1201 Eastlake Avenue East  
; CITY: Seattle  
; STATE: WA  
; COUNTRY: USA  
; ZIP: 98102  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Diskette  
; COMPUTER: IBM Compatible  
; OPERATING SYSTEM: DOS  
; SOFTWARE: FastSeq for Windows Version 2.0  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/991.890  
; FILING DATE:  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/033.003  
; FILING DATE: December 16, 1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Sawislak, Deborah A.  
; REGISTRATION NUMBER: 37,438  
; REFERENCE/DOCKET NUMBER: 96-41  
; TELEPHONE: 206-442-6672  
; TELEFAX: 206-442-6678  
; TELEX:  
; INFORMATION FOR SEQ ID NO: 1:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 420 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: double  
; TOPOLOGY: linear  
; MOLECULE TYPE: cDNA  
; FEATURE:  
; NAME/KEY: Coding Sequence

```
; LOCATION: 1...417
; OTHER INFORMATION:
US-08-991-890-1

Query Match 100.0% Score 342; DB 3; Length 420;
Best Local Similarity 100.0% Pred. No. 1.7e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTGTCATATTGCCCATG 60
DB 76 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTGTCATATTGCCCATG 135
QY 61 CCTGAGAAGACATTACCAACACCACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 120
DB 136 CCTGAGAAGACATTACCAACACCACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 195
QY 121 GAAATGGTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 180
DB 196 GAAATGGTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 255
QY 181 ATTCTTAATTTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 240
DB 256 ATTCTTAATTTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 315
QY 241 AAATTAATACTTTCCCGCAAAAGAGAGTGGAGTACAGATTTCAGTAAATTAATGTACA 342
DB 316 AAATTAATACTTTCCCGCAAAAGAGAGTGGAGTACAGATTTCAGTAAATTAATGTACA 375
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342
DB 376 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 417

RESULT 2
US-08-991-890-3
; Sequence 3, Application US/08991890
; Patent No. 6114307
; GENERAL INFORMATION:
; APPLICANT: Jaspers, Stephen R.
; APPLICANT: Sprugel, Katherine H.
; APPLICANT: Ren, Hong Ping
; APPLICANT: Humes, Jacqueline M.
; APPLICANT: Hoffman, Ross C.
; APPLICANT: Conklin, Darrell C.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: STIMULATING PANCREATIC ISLET CELL REGENERATION
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: ZymoGenetics, Inc.
; STREET: 1201 Eastlake Avenue East
; CITY: Seattle
; STATE: WA
; COUNTRY: USA
; ZIP: 98102
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; FILING DATE:
; FILING NUMBER: US/08/991,890
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,003
; FILING DATE: December 16, 1996
; ATTORNEY/AGENT INFORMATION:
; NAME: Sawislak, Deborah A
; REGISTRATION NUMBER: 37,438
; REFERENCE/DOCKET NUMBER: 96-41
; TELEPHONE: 206-442-6672
; TELEFAX: 206-442-6678

; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 480 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: Coding Sequence
; LOCATION: 1...477
; OTHER INFORMATION:
US-08-991-890-3

Query Match 100.0% Score 342; DB 3; Length 480;
Best Local Similarity 100.0% Pred. No. 1.8e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTCTCTCATATTGCCCATG 60
DB 136 GCAGAGCTGAGGGGATGTGGTCCCGATTGGAAAAACACTTGTCTCTCATATTGCCCATG 195
QY 61 CCTGAGAAGACATTACCAACACCACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 120
DB 196 CCTGAGAAGACATTACCAACACCACCCAGAGGGTGGTCTGGAATCTGGACGTCCCAAA 255
QY 121 GAAATGGTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 180
DB 256 GAAATGGTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 315
QY 181 ATTCTTAATTTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 240
DB 316 ATTCTTAATTTGTCAACCTCCACACACAAAGATGGACAGCCITAGGTAGACATCAGATTTC 375
QY 241 AAATTAATACTTTCCCGCAAAAGAGAGTGGAGTACAGATTTCAGTAAATTAATGTACA 300
DB 376 AAATTAATACTTTCCCGCAAAAGAGAGTGGAGTACAGATTTCAGTAAATTAATGTACA 435
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342
DB 436 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 477

RESULT 3
US-08-482-842B-3
; Sequence 3, Application US/08482842B
; Patent No. 5910480
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahmet
; APPLICANT: BELLET, Dominique
; APPLICANT: CHASSIN, Dorine
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN,
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF THIS PROTEIN AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION CONTAINING SUCH, DNA CODING FOR
; TITLE OF INVENTION: SAID PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,842B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
```

```
; APPLICATION NUMBER: FR 94-07191
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 017753-058
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 107..523
; US-08-482-842B-3

Query Match      100.0%; Score 342; DB 2; Length 615;
Best Local Similarity 100.0%; Pred. No. 2e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 60
DB 182 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 241
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 120
DB 242 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 301
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAAGCTTAGGTACGACATCAGAATTC 180
DB 302 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAAGCTTAGGTACGACATCAGAATTC 361
QY 181 ATTCCTAATTTGTCACCAAGCTGGAACACCACTGCTGAAGGCGACCCATCATTTGAAG 240
DB 362 ATTCCTAATTTGTCACCAAGCTGGAACACCACTGCTGAAGGCGACCCATCATTTGAAG 421
QY 241 AAAATATATCTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATCCATTCGTTGT 300
DB 422 AAAATATATCTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATCCATTCGTTGT 481
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342
DB 482 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 523

RESULT 4
US-09-174-465D-4
; Sequence 4, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dominique
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; CURRENT FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
; US-09-599-564A-4

Query Match      100.0%; Score 342; DB 4; Length 934;
Best Local Similarity 100.0%; Pred. No. 2.4e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 60
DB 501 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 120
DB 501 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 120
```

```
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
; US-09-174-465D-4

Query Match      100.0%; Score 342; DB 3; Length 934;
Best Local Similarity 100.0%; Pred. No. 2.4e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 60
DB 501 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 120
DB 561 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 620
QY 121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAAGCTTAGGTACGACATCAGAATTC 180
DB 621 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAAGCTTAGGTACGACATCAGAATTC 680
QY 181 ATTCCTAATTTGTCACCAAGCTGGAACACCACTGCTGAAGGCGACCCATCATTTGAAG 240
DB 681 ATTCCTAATTTGTCACCAAGCTGGAACACCACTGCTGAAGGCGACCCATCATTTGAAG 740
QY 241 AAAATATATCTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATCCATTCGTTGT 300
DB 741 AAAATATATCTTTCCCGGAAAAGAGAGTGGACGTACAGATTTGATCCATTCGTTGT 800
QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 342
DB 801 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATATGTACA 842

RESULT 5
US-09-599-564A-4
; Sequence 4, Application US/09599564A
; Patent No. 6362318
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Ahment
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-127
; CURRENT APPLICATION NUMBER: US/09/599,564A
; CURRENT FILING DATE: 2000-06-23
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 4
; LENGTH: 934
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
; US-09-599-564A-4

Query Match      100.0%; Score 342; DB 4; Length 934;
Best Local Similarity 100.0%; Pred. No. 2.4e-101;
Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 60
DB 501 GCAGAGCTGAGGGGATGTCCTCCCGATTTGGAAACACTTGTCTGCATATGCCCCATG 560
QY 61 CCTGAGAAGACATTCACCAACACCCAGGAGGTGGCTGGTGAATCTGGAGCTCCCAAA 120
```

```
Db 561 CCTGAGAGACATTCACCAACCCAGGAGGTGCTGTAATCTGGAGTCCCAAA 620
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 621 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 680
Qy 181 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTGAAG 240
Db 681 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTGAAG 740
Qy 241 AAAATAATACCTTTCCCGAAAAAGAGAAGTGGAGCTCAGATTTGATCCATCTGTTGT 300
Db 741 AAAATAATACCTTTCCCGAAAAAGAGAAGTGGAGCTCAGATTTGATCCATCTGTTGT 800
Qy 301 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATTAATGTACA 342
Db 801 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATTAATGTACA 842

RESULT 6
US-08-482-842B-1
; Sequence 1, Application US/08482842B
; Patent No. 5910480
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahmet
; APPLICANT: CHASSIN, Dordine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN,
; TITLE OF INVENTION: PROCESS FOR THE PREPARATION OF THIS PROTEIN AND
; TITLE OF INVENTION: PHARMACEUTICAL COMPOSITION CONTAINING SUCH, DNA CODING FOR
; TITLE OF INVENTION: SAID PROTEIN
; NUMBER OF SEQUENCES: 7
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Burns, Doane, Swecker & Mathis
; STREET: P.O. Box 1404
; CITY: Alexandria
; STATE: Virginia
; COUNTRY: United States
; ZIP: 22313-1404
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/482,842B
; FILING DATE: 07-JUN-1995
; CLASSIFICATION: 514
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: FR 94-07191
; FILING DATE: 13-JUN-1994
; ATTORNEY/AGENT INFORMATION:
; NAME: McGowan, Malcolm K.
; REGISTRATION NUMBER: 39,300
; REFERENCE/DOCKET NUMBER: 017753-058
; TELEPHONE: (703) 836-6620
; TELEFAX: (703) 836-2021
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 615 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: DNA (genomic)
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..615
; US-08-482-842B-1

Query Match 99.5%; Score 340.4; DB 2; Length 615;
Best Local Similarity 99.7%; Pred. No. 6.6e-101;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Db 561 CCTGAGAGACATTCACCAACCCAGGAGGTGCTGTAATCTGGAGTCCCAAA 620
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 621 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 680
Qy 181 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTGAAG 240
Db 681 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTGAAG 740
Qy 241 AAAATAATACCTTTCCCGAAAAAGAGAAGTGGAGCTCAGATTTGATCCATCTGTTGT 300
Db 741 AAAATAATACCTTTCCCGAAAAAGAGAAGTGGAGCTCAGATTTGATCCATCTGTTGT 800
Qy 301 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATTAATGTACA 342
Db 801 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATTAATGTACA 842

RESULT 7
US-09-174-465D-1
; Sequence 1, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahmet
; APPLICANT: CHASSIN, Dordine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; CURRENT FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 615
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(615)
; OTHER INFORMATION: Description of Unknown Organism: EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like Peptide
; US-09-174-465D-1

Query Match 99.5%; Score 340.4; DB 3; Length 615;
Best Local Similarity 99.7%; Pred. No. 6.6e-101;
Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
Qy 1 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAAACACTTCTGCTATATTGCCCATG 60
Db 181 GCAGAGCTGAGGGGATGTGTCCTCCCGATTTGGAAAACACTTCTGCTATATTGCCCATG 240
Qy 61 CCTGAGAAGACATTCACCAACCCAGGAGGTGCTGCTGGAATCTGGACGTCCCAAA 120
Db 241 CCTGAGAAGACATTCACCAACCCAGGAGGTGCTGCTGGAATCTGGACGTCCCAAA 300
Qy 121 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 180
Db 301 GAAATGGTCTCAACCTCCACAACAAGATGGACAAGCCCTTAGGTACGACATCAGAAATTC 360
Qy 181 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTGAAG 240
Db 361 ATTCCTAATTTGTCCACAGAGCTGAAGAAACCACTGCTGTAAGGGCAGCCATCATTGAAG 420
Qy 241 AAAATAATACCTTTCCCGAAAAAGAGAAGTGGAGCTCAGATTTGATCCATCTGTTGT 300
Db 421 AAAATAATACCTTTCCCGAAAAAGAGAAGTGGAGCTCAGATTTGATCCATCTGTTGT 480
Qy 301 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATTAATGTACA 342
Db 481 GAAGTAATTTGTGAGGTGAACCTTCAGTTAAATTAATGTACA 522
```

Qy	1	GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACCTTGCTGTCTATATTGCCCATG	60
Db	181	GCAGAGCTGAGGGATGTGGTCCCGGATTTGGAAACACCTTGCTGTCTATATTGCCCATG	240
Qy	61	CCTGAGAAGACATTCACCAACACCCAGAGGGTGGCTGCTGGAATCTGGACGCTCCCAA	120
Db	241	CCTGAGAAGACATTCACCAACACCCAGAGGGTGGCTGCTGGAATCTGGACGCTCCCAA	300
Qy	121	GAATCGGTGTCACCTCCCAACAAAGATGGACAAGCCCTTAGCTACGCATCAGAAATTC	180
Db	301	GAATCGGTGTCACCTCCCAAGAACAAAGATGGACAAGCCCTTAGCTACGCATCAGAAATTC	360
Qy	181	ATTCTTAATTTGTCACAGAGTGAAGAAACCACTGTCTGAAGGCGACCCATCATTGAAG	240
Db	361	ATTCTTAATTTGTCACAGAGTGAAGAAACCACTGTCTGAAGGCGACCCATCATTGAAG	420
Qy	241	AAAATAATACTCTCCCGCAAAAAGAGAAGTGGACGCTCACAGATTTGATCCATTCGTGTGT	300
Db	421	AAAATAATACTCTCCCGCAAAAAGAGAAGTGGACGCTCACAGATTTGATCCATTCGTGTGT	480
Qy	301	GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA	342
Db	481	GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA	522

```

1 RESULT 10
2 US-09-599-564A-9
3 ; Sequence 9, Application US/09599564A
4 ; Patent No. 6362318
5 ; GENERAL INFORMATION:
6 ; APPLICANT: KOMAN, Ahment
7 ; APPLICANT: CHASSIN, Dorine
8 ; APPLICANT: BELLET, Dominique
9 ; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
10 ; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
11 ; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
12 ; TITLE OF INVENTION: PROTEIN
13 ; FILE REFERENCE: 017753-127
14 ; CURRENT APPLICATION NUMBER: US/09/599,564A
15 ; CURRENT FILING DATE: 2000-06-23
16 ; PRIOR APPLICATION NUMBER: 09/174,465
17 ; PRIOR FILING DATE: 1998-10-19
18 ; PRIOR APPLICATION NUMBER: US 08/482,842
19 ; PRIOR FILING DATE: 1995-06-07
20 ; NUMBER OF SEQ ID NOS: 16
21 ; SOFTWARE: PatentIn Ver. 2.0
22 ; SEQ ID NO 9
23 ; LENGTH: 153
24 ; TYPE: DNA
25 ; ORGANISM: Unknown

```

```
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(153)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-599-564A-9

Query Match
Best Local Similarity 44.7%; Score 153; DB 4; Length 153;
Matches 153; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 100 CTGAATCTGGACGTCCTCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGCAAGCC 159
Db 1 CTGAATCTGGACGTCCTCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGCAAGCC 60

Qy 160 TTAGGTAGACATCAGAAATCAATTCCTAAATTTGTCAACGAGCTGAAGAAACCACTGTCT 219
Db 61 TTAGGTAGACATCAGAAATCAATTCCTAAATTTGTCAACGAGCTGAAGAAACCACTGTCT 120

Qy 220 GAAGGGCAGCCATCATTTGAAGAAATAATACTT 252
Db 121 GAAGGGCAGCCATCATTTGAAGAAATAATACTT 153

RESULT 11
US-09-174-465D-7
; Sequence 7, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dorine
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; PRIOR FILING DATE: 1998-10-19
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 123
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(123)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-174-465D-7

Query Match
Best Local Similarity 28.9%; Score 99; DB 3; Length 123;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAAACACACTTGTCTCATATTGCCCATG 60
Db 25 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAAACACACTTGTCTCATATTGCCCATG 84

Qy 61 CCTGAGAAGACATTCACACACCCAGGAGGGTGGCTG 99
Db 85 CCTGAGAAGACATTCACACACCCAGGAGGGTGGCTG 123

RESULT 12
US-09-599-564A-7
; Sequence 7, Application US/09599564A
; Patent No. 6362318
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
```

```
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; TITLE OF INVENTION: PROTEIN
; FILE REFERENCE: 017753-127
; CURRENT APPLICATION NUMBER: US/09/599,564A
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: 09/174,465
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 7
; LENGTH: 123
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(123)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-599-564A-7

Query Match
Best Local Similarity 28.9%; Score 99; DB 4; Length 123;
Matches 99; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAAACACACTTGTCTCATATTGCCCATG 60
Db 25 GCAGAGCTGAGGGATGTGTCCTCCGATTTGGAAACACACTTGTCTCATATTGCCCATG 84

Qy 61 CCTGAGAAGACATTCACACACCCAGGAGGGTGGCTG 99
Db 85 CCTGAGAAGACATTCACACACCCAGGAGGGTGGCTG 123

RESULT 13
US-09-174-465D-11
; Sequence 11, Application US/09174465D
; Patent No. 6180364
; GENERAL INFORMATION:
; APPLICANT: KOMAN, Ahment
; APPLICANT: CHASSIN, Dorine
; APPLICANT: BELLET, Dominique
; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE
; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL
; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID
; FILE REFERENCE: 017753-103
; CURRENT APPLICATION NUMBER: US/09/174,465D
; PRIOR FILING DATE: 1998-10-19
; PRIOR APPLICATION NUMBER: US 08/482,842
; PRIOR FILING DATE: 1995-06-07
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
; LENGTH: 93
; TYPE: DNA
; ORGANISM: Unknown
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(93)
; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early
; OTHER INFORMATION: Placenta Insulin-Like peptide
US-09-174-465D-11

Query Match
Best Local Similarity 26.3%; Score 90; DB 3; Length 93;
Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
```



QY 253 TCCCGCAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 312  
 Db 1 TCCCGCAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 60  
 QY 313 GACGATGGAACCTTCAGTTAAATTATGTACA 342  
 Db 61 GACGATGGAACCTTCAGTTAAATTATGTACA 90

RESULT 14  
 US-09-599-564A-11  
 ; Sequence 11, Application US/09599564A  
 ; Patent No. 6362318  
 ; GENERAL INFORMATION:  
 ; APPLICANT: KOMAN, Alment  
 ; APPLICANT: CHASSIN, Dorine  
 ; APPLICANT: BELLET, Dominique  
 ; TITLE OF INVENTION: NEW PROTEIN CALLED EPIL/PLACENTIN, PROCESS FOR THE  
 ; PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
 ; TITLE OF INVENTION: PREPARATION OF THIS PROTEIN AND PHARMACEUTICAL  
 ; TITLE OF INVENTION: COMPOSITION CONTAINING SUCH, DNA CODING FOR SAID  
 ; TITLE OF INVENTION: PROTEIN  
 ; FILE REFERENCE: 017753-127  
 ; CURRENT APPLICATION NUMBER: US/09/599,564A  
 ; CURRENT FILING DATE: 2000-06-23  
 ; PRIOR APPLICATION NUMBER: 09/174,465  
 ; PRIOR FILING DATE: 1998-10-19  
 ; PRIOR APPLICATION NUMBER: US 08/482,842  
 ; PRIOR FILING DATE: 1995-06-07  
 ; NUMBER OF SEQ ID NOS: 16  
 ; SOFTWARE: PatentIn Ver. 2.0  
 ; SEQ ID NO 11  
 ; LENGTH: 93  
 ; TYPE: DNA  
 ; ORGANISM: Unknown  
 ; FEATURE:  
 ; NAME/KEY: CDS  
 ; LOCATION: (1)..(93)  
 ; OTHER INFORMATION: Description of Unknown Organism:EPIL - Early  
 ; OTHER INFORMATION: Placenta Insulin-Like peptide  
 US-09-599-564A-11

Query Match 26.3%; Score 90; DB 4; Length 93;  
 Best Local Similarity 100.0%; Pred. No. 4.8e-20;  
 Matches 90; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 253 TCCCGCAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 312  
 Db 1 TCCCGCAAAAGAGAGTGGAGCTCACAGATTGTATCCATTCTGTTGTGAAGTAATTGT 60  
 QY 313 GACGATGGAACCTTCAGTTAAATTATGTACA 342  
 Db 61 GACGATGGAACCTTCAGTTAAATTATGTACA 90

RESULT 15  
 5464756-17  
 ; Patent No. 5464756  
 ; APPLICANT: HENNER, DENNIS J.; VANDLEN, RICHARD L.; WILKINS,  
 ; JAMES A.; YANSURA, DANIEL G.  
 ; TITLE OF INVENTION: PROCESS AND COMPOSITIONS FOR THE  
 ; ISOLATION HUMAN RELAXIN  
 ; NUMBER OF SEQUENCES: 42  
 ; CURRENT APPLICATION DATA:  
 ; APPLICATION NUMBER: US/07/908,766  
 ; FILING DATE: 01-JUL-1992  
 ; PRIOR APPLICATION DATA:  
 ; APPLICATION NUMBER: 347,550  
 ; FILING DATE: 04-MAY-1989  
 ; SEQ ID NO:17;  
 ; LENGTH: 489  
 5464756-17

Query Match 21.0%; Score 71.8; DB 6; Length 489;

Best Local Similarity 66.5%; Pred. No. 7.9e-14;  
 Matches 103; Conservative 0; Mismatches 52; Indels 0; Gaps 0;  
 QY 82 ACCCCAGGAGGGTGGCTGCTGGAATCTGGACGTCCCAAGAAATGGTGTCACCTCCAC 141  
 Db 103 AGCCAGGAGATGCTCTCTCAGACACCTAGACCAGTGGCAGAAATTTGGCCATCCTC 162  
 QY 142 AACAAAGATGGACAAAGCCCTTAGGTAGCATCAGAAATTCATTCCTAAATTTGTCAC 201  
 Db 163 AACAAAGATACAGAAACCATAAATATGATGTGCAAAATTTGTTGCTAAATTTGCC 222  
 QY 202 CTGAAGAAACCCACTGTCTGAAGGCGACCATCAT 236  
 Db 223 CTGAAGTTAACCCCTGTCTGAGATGCGACGCAAT 257

Search completed: October 17, 2003, 02:46:30  
 Job time : 54 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2003, 02:21:42 ; Search time 183 Seconds  
(without alignments)

4907.846 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GCAGAGCTGAGGGGATGTGG.....CTTCAGTTAAATTATGTACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 1750203 seqs, 1313063994 residues

Total number of hits satisfying chosen parameters: 3500406

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

Published\_Applications\_NA.\*  
1: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
2: /cgn2\_6/ptodata/2/pubpna/PCT\_NEW\_PUB.seq.\*  
3: /cgn2\_6/ptodata/2/pubpna/US06\_NEW\_PUB.seq.\*  
4: /cgn2\_6/ptodata/2/pubpna/US06\_PUBCOMB.seq.\*  
5: /cgn2\_6/ptodata/2/pubpna/US07\_NEW\_PUB.seq.\*  
6: /cgn2\_6/ptodata/2/pubpna/US07\_PUBCOMB.seq.\*  
7: /cgn2\_6/ptodata/2/pubpna/US08\_NEW\_PUB.seq.\*  
8: /cgn2\_6/ptodata/2/pubpna/US08\_PUBCOMB.seq.\*  
9: /cgn2\_6/ptodata/2/pubpna/US09A\_PUBCOMB.seq.\*  
10: /cgn2\_6/ptodata/2/pubpna/US09B\_PUBCOMB.seq.\*  
11: /cgn2\_6/ptodata/2/pubpna/US09C\_PUBCOMB.seq.\*  
12: /cgn2\_6/ptodata/2/pubpna/US09\_NEW\_PUB.seq.\*  
13: /cgn2\_6/ptodata/2/pubpna/US10A\_PUBCOMB.seq.\*  
14: /cgn2\_6/ptodata/2/pubpna/US10B\_PUBCOMB.seq.\*  
15: /cgn2\_6/ptodata/2/pubpna/US10\_NEW\_PUB.seq.\*  
16: /cgn2\_6/ptodata/2/pubpna/US60\_NEW\_PUB.seq.\*  
17: /cgn2\_6/ptodata/2/pubpna/US60\_PUBCOMB.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	69	20.2	231	9	US-09-759-143-465 Sequence 465, App
C 2	69	20.2	231	9	US-09-780-669-465 Sequence 465, App
C 3	69	20.2	231	9	US-09-822-827-465 Sequence 465, App
C 4	69	20.2	231	10	US-09-895-793-465 Sequence 465, App
C 5	69	20.2	231	10	US-09-895-814-465 Sequence 465, App
C 6	69	20.2	231	12	US-10-144-678A-465 Sequence 465, App
C 7	69	20.2	231	12	US-10-294-025-465 Sequence 465, App
C 8	69	20.2	231	13	US-10-012-896-465 Sequence 465, App
C 9	69	20.2	231	14	US-10-010-940-465 Sequence 465, App
10	68.6	20.1	558	14	US-10-205-823-348 Sequence 348, App
11	60	17.5	60	12	US-09-908-975-31961 Sequence 31961, A
12	38.6	11.3	2574	9	US-09-735-103-2 Sequence 2, Appli
13	38.6	11.3	2574	13	US-10-045-428A-2 Sequence 2, Appli
C 14	35	10.2	116592	10	US-09-818-512-3 Sequence 3, Appli
C 15	35	10.2	116592	12	US-10-354-065-3 Sequence 3, Appli
16	33	9.6	528	13	US-10-027-632-298520 Sequence 298520, Sequence 111160, Sequence 6894, Ap

Sequence 2481, Ap  
Sequence 246198,  
Sequence 382, App  
Sequence 30248, A  
Sequence 30249, A  
Sequence 30250, A  
Sequence 313, App  
Sequence 313, App  
Sequence 312, App  
Sequence 93, Appl  
Sequence 35, Appl  
Sequence 314, Appl  
Sequence 6, Appl  
Sequence 3396, Ap  
Sequence 223414,  
Sequence 134, App  
Sequence 3, Appl  
Sequence 9632, Ap  
Sequence 3869, Ap  
Sequence 185112,  
Sequence 187, App  
Sequence 215, App  
Sequence 8229, Ap  
Sequence 164, App

## ALIGNMENTS

### RESULT 1

US-09-759-143-465/c  
; Sequence 465, Application US/09759143  
; Patent No. US2002002248A1

### GENERAL INFORMATION:

APPLICANT: Xu, Jiangchun  
APPLICANT: Dillon, Davin C.  
APPLICANT: Mitcham, Jennifer L.  
APPLICANT: Harlocker, Susan L.  
APPLICANT: Jiang, Yuqi  
APPLICANT: Henderson, Robert A.  
APPLICANT: Kalos, Michael D.  
APPLICANT: Fanger, Gary R.  
APPLICANT: Retter, Marc W.  
APPLICANT: Stolk, John A.  
APPLICANT: Day, Craig H.  
APPLICANT: Vedvick, Thomas S.  
APPLICANT: Carter, Darrick  
APPLICANT: Li, Samuel  
APPLICANT: Wang, Aijun  
APPLICANT: Skeiky, Yasir A.W.  
APPLICANT: Hepler, William

TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND  
FILE REFERENCE: 210121.427C23  
CURRENT APPLICATION NUMBER: US/09759,143  
CURRENT FILING DATE: 2001-01-12  
NUMBER OF SEQ ID NOS: 934  
SOFTWARE: FastSeq for Windows Version 3.0  
SEQ ID NO 465  
LENGTH: 231  
TYPE: DNA  
ORGANISM: Homo sapiens  
US-09-759-143-465

Query Match 20.2%; Score 69; DB 9; Length 231;  
Best Local Similarity 74.4%; Pred. No. 1.2e-12;  
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

## RESULT 5

```
US-09-895-814-465/c
; Sequence 465, Application US/09895814
; Publication No. US20020193296A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C26
; CURRENT APPLICATION NUMBER: US/09/895,814
; CURRENT FILING DATE: 2001-05-29
; NUMBER OF SEQ ID NOS: 990
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-895-814-465

Query Match
Best Local Similarity 74.4%; Score 69; DB 10; Length 231;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAGCCTTAGGTACGACATCAGAATT 179
DB 137 AAAAAATTGGCCATCCTTCATCAACAAAGATACAGAAACCATAAATATGATGTCAGAATT 78

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGACCACTGTCTGAAGGGCAGCCATCAIT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCCCTGTCTGAGATGCAGCCAGCATT 21

RESULT 6
US-10-144-678A-465/c
; Sequence 465, Application US/10144678A
; Publication No. US20030157089A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yuqiu
; APPLICANT: Henderson, Robert A.
; APPLICANT: Kalos, Michael D.
; APPLICANT: Fanger, Gary R.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedvick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skeiky, Yasir A. W.

US-09-895-814-465/c
; Sequence 465, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-294-025-465

Query Match
Best Local Similarity 74.4%; Score 69; DB 12; Length 231;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAGCCTTAGGTACGACATCAGAATT 179
DB 137 AAAAAATTGGCCATCCTTCATCAACAAAGATACAGAAACCATAAATATGATGTCAGAATT 78

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGACCACTGTCTGAAGGGCAGCCATCAIT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCCCTGTCTGAGATGCAGCCAGCATT 21

RESULT 7
US-10-294-025-465/c
; Sequence 465, Application US/10294025
; Publication No. US20030185830A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Stolk, John A.
; APPLICANT: Kalos, Michael D.
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C29
; CURRENT APPLICATION NUMBER: US/10/294,025
; CURRENT FILING DATE: 2002-11-12
; NUMBER OF SEQ ID NOS: 1038
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-294-025-465

Query Match
Best Local Similarity 74.4%; Score 69; DB 12; Length 231;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAGCCTTAGGTACGACATCAGAATT 179
DB 137 AAAAAATTGGCCATCCTTCATCAACAAAGATACAGAAACCATAAATATGATGTCAGAATT 78

QY 180 CATTCTTAATTTGTCACAGAGCTGAAGACCACTGTCTGAAGGGCAGCCATCAIT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCCCTGTCTGAGATGCAGCCAGCATT 21

RESULT 8
US-10-012-896-465/c
; Sequence 465, Application US/10012896
; Publication No. US20020183251A1
; GENERAL INFORMATION:
```

```

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Jiang, Yudi
; APPLICANT: Kalos, Michael D.
; APPLICANT: Retter, Marc W.
; APPLICANT: Stolk, John A.
; APPLICANT: Day, Craig H.
; APPLICANT: Vedwick, Thomas S.
; APPLICANT: Carter, Darrick
; APPLICANT: Li, Samuel X.
; APPLICANT: Wang, Aijun
; APPLICANT: Skelky, Yasir A.W.
; APPLICANT: Hepler, William T.
; APPLICANT: Henderson, Robert A.
; APPLICANT: Hural, John
; APPLICANT: McNeill, Patricia D.
; APPLICANT: Houghton, Raymond L.
; APPLICANT: Vinals de Bassols, Carlota
; APPLICANT: Foy, Teresa
; APPLICANT: Fanger, Gary R.
; APPLICANT: Wantanabe, Yoshihiro
; APPLICANT: Meagher, Madeleine Joy
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427C27
; CURRENT APPLICATION NUMBER: US/10/012,896
; CURRENT FILING DATE: 2001-12-10
; NUMBER OF SEQ ID NOS: 1011
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-465

Query Match      20.2%; Score 69; DB 13; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-010-940-465

Query Match      20.2%; Score 69; DB 14; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAAGCCTTAGGTAGCAGCATCAGAATT 179
DB 137 AAAAATTGGCCATCTTCATCAACAAGATACAGAAACCATTAATATGTGTGAGAATT 78
QY 180 CATTCCTAATTTGTCAACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCTGTCTGTGAGATCGCCAGCAT 21

RESULT 9
US-10-010-940-465/c
; Sequence 465, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-465

Query Match      20.2%; Score 69; DB 13; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAAGCCTTAGGTAGCAGCATCAGAATT 179
DB 137 AAAAATTGGCCATCTTCATCAACAAGATACAGAAACCATTAATATGTGTGAGAATT 78
QY 180 CATTCCTAATTTGTCAACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCTGTCTGTGAGATCGCCAGCAT 21

RESULT 9
US-10-010-940-465/c
; Sequence 465, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-465

Query Match      20.2%; Score 69; DB 13; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAAGCCTTAGGTAGCAGCATCAGAATT 179
DB 137 AAAAATTGGCCATCTTCATCAACAAGATACAGAAACCATTAATATGTGTGAGAATT 78
QY 180 CATTCCTAATTTGTCAACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCTGTCTGTGAGATCGCCAGCAT 21

RESULT 9
US-10-010-940-465/c
; Sequence 465, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-465

Query Match      20.2%; Score 69; DB 13; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAAGCCTTAGGTAGCAGCATCAGAATT 179
DB 137 AAAAATTGGCCATCTTCATCAACAAGATACAGAAACCATTAATATGTGTGAGAATT 78
QY 180 CATTCCTAATTTGTCAACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCTGTCTGTGAGATCGCCAGCAT 21

RESULT 9
US-10-010-940-465/c
; Sequence 465, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-465

Query Match      20.2%; Score 69; DB 13; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAAGCCTTAGGTAGCAGCATCAGAATT 179
DB 137 AAAAATTGGCCATCTTCATCAACAAGATACAGAAACCATTAATATGTGTGAGAATT 78
QY 180 CATTCCTAATTTGTCAACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCTGTCTGTGAGATCGCCAGCAT 21

RESULT 9
US-10-010-940-465/c
; Sequence 465, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
; FILE REFERENCE: 210121.427D3
; CURRENT APPLICATION NUMBER: US/10/010,940
; CURRENT FILING DATE: 2001-12-05
; NUMBER OF SEQ ID NOS: 575
; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 465
; LENGTH: 231
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-012-896-465

Query Match      20.2%; Score 69; DB 13; Length 231;
Best Local Similarity 74.4%; Pred. No. 1.2e-12;
Matches 87; Conservative 0; Mismatches 30; Indels 0; Gaps 0;

QY 120 AGAATGGTGTCAACCTCCACAAAGATGGACAAGCCTTAGGTAGCAGCATCAGAATT 179
DB 137 AAAAATTGGCCATCTTCATCAACAAGATACAGAAACCATTAATATGTGTGAGAATT 78
QY 180 CATTCCTAATTTGTCAACAGAGCTGAAGAAACCACTGTCTGAAGGGCAGCCATCAT 236
DB 77 TGTGCTAATTTGCCACAGGAGCTGAAGTTAACCTGTCTGTGAGATCGCCAGCAT 21

RESULT 9
US-10-010-940-465/c
; Sequence 465, Application US/10010940
; Publication No. US2003008062A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Dillon, Davin C.
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan Louise
; APPLICANT: Jiang Yuqi
; APPLICANT: Reed, Steven G.
; APPLICANT: Kalos, Michael
; APPLICANT: Fanger, Gary
; APPLICANT: Retter, Mark
; APPLICANT: Solk, John
; APPLICANT: Day, Craig
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THERAPY AND
; TITLE OF INVENTION: DIAGNOSIS OF PROSTATE CANCER
;
```

Db 292 CTGAAGCAGCCCTATCTGAGAGGCAACCATCATT 326

## RESULT 11

US-09-908-975-31961  
; Sequence 31961, Application US/09908975  
; Publication No. US20030165843A1  
; GENERAL INFORMATION:  
; APPLICANT: SHOSHAN, Avi  
; APPLICANT: WASSERMAN, Alon  
; APPLICANT: MINTZ, Eli  
; APPLICANT: MINTZ, Liat  
; APPLICANT: FAIGLER, Simchon  
; TITLE OF INVENTION: OLIGONUCLEOTIDE LIBRARY FOR DETECTING RNA TRANSCRIPTS AND SPLICING  
; TITLE OF INVENTION: THAT POPULATE A TRANSCRIPTOME  
; FILE REFERENCE: 36888-0005  
; CURRENT APPLICATION NUMBER: US/09/908,975  
; CURRENT FILING DATE: 2001-07-20  
; PRIOR APPLICATION NUMBER: US 60/287,724  
; PRIOR FILING DATE: 2001-05-02  
; PRIOR APPLICATION NUMBER: US 60/221,607  
; PRIOR FILING DATE: 2000-07-28  
; NUMBER OF SEQ ID NOS: 32337  
; SOFTWARE: PatentIn version 3.0  
; SEQ ID NO 31961  
; LENGTH: 60  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-908-975-31961

Query Match 17.5%; Score 60; DB 12; Length 60;  
Best Local Similarity 100.0%; Pred. No. 5.8e-10;  
Matches 60; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 222 AGGGCAGCCATCATTTGAAGAAATATCTTCCGCAAAAGAGAGAGTGGCGTCACAG 281  
DB 1 AGGGCAGCCATCATTTGAAGAAATATCTTCCGCAAAAGAGAGAGTGGCGTCACAG 60

## RESULT 12

US-09-735-103-2  
; Sequence 2, Application US/09735103  
; Patent No. US20010005589A1  
; GENERAL INFORMATION:  
; APPLICANT: Mano, Hiroyuki  
; APPLICANT: Sakata, Tsuneaki  
; APPLICANT: Hasegawa, Mamoru  
; TITLE OF INVENTION: Promoter  
; FILE REFERENCE: 50026/011002  
; CURRENT APPLICATION NUMBER: US/09/735,103  
; CURRENT FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 09/142,529  
; PRIOR FILING DATE: 1998-09-09  
; PRIOR APPLICATION NUMBER: PCT/JP97/00741  
; PRIOR FILING DATE: 1997-03-10  
; PRIOR APPLICATION NUMBER: JP 8/54294  
; PRIOR FILING DATE: 1996-03-12  
; NUMBER OF SEQ ID NOS: 9  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2574  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-09-735-103-2

Query Match 11.3%; Score 38.6; DB 9; Length 2574;  
Best Local Similarity 52.9%; Pred. No. 0.058;  
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 54 CCCATGCGCTGAGAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACG 113  
DB 102 CCAGAAGACCGAGATGAATTTCAACACTATCTTAGAAGAGATTTCTATTAAAGGTCCTCA 161

QY 114 TCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTACGACATC 173  
DB 162 GCAGAAAAGAGACATCATCTTTAAACTACAAGAGAGACTTTGTGTACTTCCAAAATC 221  
QY 174 AGAATTCATTCTTAATTTGTCCACGAGAGCTGGAAGAAA 210  
DB 222 CGTGTTCAGCTACTATGAGGGTCGAGCGGAGAGAAA 258

## RESULT 13

US-10-045-428A-2  
; Sequence 2, Application US/10045428A  
; Publication No. US20020115845A1  
; GENERAL INFORMATION:  
; APPLICANT: Mano, Hiroyuki  
; APPLICANT: Sakata, Tsuneaki  
; APPLICANT: Hasegawa, Mamoru  
; APPLICANT: Tabata, Toshiaki  
; TITLE OF INVENTION: Promoter  
; FILE REFERENCE: 50026/011003  
; CURRENT APPLICATION NUMBER: US/10/045,428A  
; CURRENT FILING DATE: 2002-04-15  
; PRIOR APPLICATION NUMBER: 09/735,103  
; PRIOR FILING DATE: 2000-12-12  
; PRIOR APPLICATION NUMBER: 09/142,529  
; PRIOR FILING DATE: 1999-08-12  
; PRIOR APPLICATION NUMBER: PCT/JP97/00741  
; PRIOR FILING DATE: 1997-03-10  
; PRIOR APPLICATION NUMBER: JP 8/54294  
; PRIOR FILING DATE: 1996-03-12  
; NUMBER OF SEQ ID NOS: 16  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 2  
; LENGTH: 2574  
; TYPE: DNA  
; ORGANISM: Mus musculus  
US-10-045-428A-2

Query Match 11.3%; Score 38.6; DB 13; Length 2574;  
Best Local Similarity 52.9%; Pred. No. 0.058;  
Matches 83; Conservative 0; Mismatches 74; Indels 0; Gaps 0;  
QY 54 CCCATGCGCTGAGAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACG 113  
DB 102 CCAGAAGACCGAGATGAATTTCAACACTATCTTAGAAGAGATTTCTATTAAAGGTCCTCA 161  
QY 114 TCCCAAGAAATGGTGTCAACCTCCCAACAAAGATGGACAGCCTTAGGTACGACATC 173  
DB 162 GCAGAAAAGAGACATCATCTTTAAACTACAAGAGAGACTTTGTGTACTTCCAAAATC 221  
QY 174 AGAATTCATTCTTAATTTGTCCACGAGAGCTGGAAGAAA 210  
DB 222 CGTGTTCAGCTACTATGAGGGTCGAGCGGAGAGAAA 258

## RESULT 14

US-09-818-512-3/c  
; Sequence 3, Application US/09818512  
; Patent No. US20020142416A1  
; GENERAL INFORMATION:  
; APPLICANT: BEASLEY, Ellen et al.  
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC  
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES  
; TITLE OF INVENTION: THEREOF  
; FILE REFERENCE: CL001192  
; CURRENT APPLICATION NUMBER: US/09/818,512  
; CURRENT FILING DATE: 2001-03-28  
; NUMBER OF SEQ ID NOS: 4  
; SOFTWARE: FastSeq for Windows Version 4.0  
; SEQ ID NO 3  
; LENGTH: 116592  
; TYPE: DNA  
; ORGANISM: Human

```
/
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
US-09-818-512-3

Query Match      10.2%; Score 35; DB 10; Length 116592;
Best Local Similarity 53.2%; Pred. No. 7;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY      101 TGGAAATCTGGACGTCCTCCAAAGAAATGGTGTCACCTCCCAACACAAAGATGGACAGCCT 160
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4383 TAGCAGATGTAGGTGGGATCGTGCGGGCCACATACACAGCACAAGACTGAAATACCT 4324

QY      161 TAGGTACGACATCAGAAATTCCTTAATTTGTCCACGAGCTGAAAGAAACCACCTGTCTG 220
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4323 AATACAGGTGATTAATAATCAATTATAGTCTATTATCTGGAATAATAATAAACTGCAAT 4264

QY      221 AAGGGCAGCCATCATTTGAA 239
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4263 CAGAGTAGCCAGAGCTGAA 4245

RESULT 15
US-10-354-065-3/c
; Sequence 3, Application US/10354065
; Publication No. US20030138837A1
; GENERAL INFORMATION:
; APPLICANT: BEASLEY, Ellen et al.
; TITLE OF INVENTION: ISOLATED HUMAN ENZYME PROTEINS, NUCLEIC
; TITLE OF INVENTION: ACID MOLECULES ENCODING HUMAN ENZYME PROTEINS, AND USES
; TITLE OF INVENTION: THEREOF
; FILE REFERENCE: CL001192DIV
; CURRENT APPLICATION NUMBER: US/10/354,065
; CURRENT FILING DATE: 2003-01-30
; NUMBER OF SEQ ID NOS: 4
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 3
; LENGTH: 116592
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(116592)
; OTHER INFORMATION: n = A,T,C or G
US-10-354-065-3

Query Match      10.2%; Score 35; DB 12; Length 116592;
Best Local Similarity 53.2%; Pred. No. 7;
Matches 74; Conservative 0; Mismatches 65; Indels 0; Gaps 0;

QY      101 TGGAAATCTGGACGTCCTCCAAAGAAATGGTGTCACCTCCCAACACAAAGATGGACAGCCT 160
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4383 TAGCAGATGTAGGTGGGATCGTGCGGGCCACATACACAGCACAAGACTGAAATACCT 4324

QY      161 TAGGTACGACATCAGAAATTCCTTAATTTGTCCACGAGCTGAAAGAAACCACCTGTCTG 220
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4323 AATACAGGTGATTAATAATCAATTATAGTCTATTATCTGGAATAATAATAAACTGCAAT 4264

QY      221 AAGGGCAGCCATCATTTGAA 239
        | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | | |
DB      4263 CAGAGTAGCCAGAGCTGAA 4245
```

Search completed: October 17, 2003, 03:26:36  
Job time : 187 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: October 17, 2003, 01:30:16 ; Search time 1428 seconds

(without alignments)  
5820.820 Million cell updates/sec

Title: US-09-518-842-1\_COPY\_76\_417

Perfect score: 342

Sequence: 1 GGAGAGCTGAGGGATGTGG.....CTTCAGTTAAATTAATGACA 342

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 22781392 seqs, 12152238056 residues

Total number of hits satisfying chosen parameters: 45562784

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estmu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_hic:\*

9: gb\_estl:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vit:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_phg:\*

27: em\_gss\_vrl:\*

28: gb\_gss1:\*

29: gb\_gss2:\*

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	342	100.0	451	9	AI088100
2	342	100.0	621	13	BX102762
3	342	100.0	816	9	AU136752
4	342	100.0	1201	9	AL550042

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

C	5	340.4	99.5	478	9	AI148843
C	6	340.4	99.5	629	12	BQ011259
C	7	331	96.8	550	9	AA77897
C	8	328.4	96.0	448	9	AI128289
C	9	321.4	94.0	523	9	AI800597
C	10	315.8	92.3	948	10	BG623890
C	11	308.4	90.2	684	14	CB549181
C	12	308.4	90.2	701	14	CB548981
C	13	307.2	89.8	424	12	BQ004108
C	14	306.8	89.7	551	14	CB549885
C	15	306.8	89.7	584	14	CB549375
C	16	306.8	89.7	646	14	CB549870
C	17	306.8	89.7	655	14	CB550346
C	18	306.8	89.7	661	14	CB550055
C	19	306.8	89.7	664	14	CB549153
C	20	306.8	89.7	665	14	CB549109
C	21	306.8	89.7	668	14	CB549740
C	22	306.8	89.7	680	14	CB550616
C	23	306.8	89.7	693	14	CB549308
C	24	306.8	89.7	726	14	CB549649
C	25	305.2	89.2	601	14	CB549945
C	26	305.2	89.2	671	14	CB549798
C	27	303.6	88.8	625	14	CB549861
C	28	297.4	87.0	569	14	CB549095
C	29	288.4	84.3	511	14	CB549723
C	30	276.6	80.9	659	14	CB549812
C	31	274.6	80.3	490	14	CB549463
C	32	265.4	77.6	414	9	AI864182
C	33	244.4	71.5	431	9	AA442679
C	34	241.2	70.5	482	14	R75651
C	35	226.4	66.2	348	12	BQ007247
C	36	212.4	62.1	427	14	R62136
C	37	211	61.7	401	14	H70182
C	38	201	58.8	329	14	R68426
C	39	199.4	58.3	414	14	H02449
C	40	191	55.8	820	10	BG197868
C	41	189.4	55.4	342	14	CB549212
C	42	183	53.5	303	12	BQ007399
C	43	181.2	53.0	705	14	CB550027
C	44	176.2	51.5	465	14	R62240
C	45	170.4	49.8	731	14	CB549758

## ALIGNMENTS

RESULT 1	AI088100/C	AI088100	451 bp	mrna	linear	EST 18-AUG-1998
LOCUS	0020a07.x1	Soares_NSF_F8_9W_OT_P_S1	Homo sapiens	CDNA clone		
DEFINITION	IMAGE:1566708 3' similar to gb:X59357 60S RIBOSOMAL PROTEIN L22 (HUMAN); mRNA sequence.					
ACCESSION	AI088100	AI088100.1	GI:3426808			
VERSION	AI088100					
KEYWORDS	EST.					
SOURCE	Homo sapiens (human)					
ORGANISM	Homo sapiens					
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homiidae; Homo.					
AUTHORS	1 (bases 1 to 451)					
TITLE	NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.					
JOURNAL	National Cancer Institute, Cancer Genome Anatomy Project (CGAP), Tumor Gene Index					
COMMENT	Unpublished					
	Contact: Robert Strausberg, Ph.D.					
	Email: cgaps@email.nih.gov					
	This clone is available royalty-free through LLNL; contact the					
	IMAGE Consortium (info@image.llnl.gov) for further information.					
	Seq primer: -40ml3 fwd. ET from Amersham					
	High quality sequence stop: 436.					
	Location/Qualifiers					
	1. .451					
	/organism="Homo sapiens"					



/mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1566708"  
 /lab\_host="DH10B"  
 /clone\_lib="Soares\_NSF\_F8\_9W\_OT\_PA\_P\_S1"  
 /note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with a modified polylinker; Site 1: Not 1; Site 2: Eco RI; Equal amounts of plasmid DNA from five normalized libraries were mixed, and ss circles were made in vitro. Following RAP purification, this DNA was used as tracer in a subtractive hybridization reaction. The driver was PCR-amplified cDNAs from pools of 5,000 clones made from the same 5 libraries. The pools consisted of the following libraries and clones: Soares NBHSF pool 1: 309384-310919, 323208-325895 Soares NB2HP pool 1: 145032-147335, 147720-148103, 148872-149255, 15002 - 150407, 151176-152327 Soares NB2HF-9W pool 1: 758280-760583, 772104-774407 Soares NB4PA pool 1: 304776-306311, 320136-322823, 326280-326663 Soares NBHOT pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."  
 pool 1: 723720-726407, 739080-740999 Subtraction by Bento Soares and M. Fatima Bonaldo."  
 BASE COUNT 115 a 95 c 97 g 144 t  
 ORIGIN

Query Match 100.0%; Score 342; DB 9; Length 451;  
 Best Local Similarity 100.0%; Pred. No. 7.7e-53;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATGCCCATG 60  
 DB 444 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATGCCCATG 385  
 QY 61 CCTGAGAGACATTCACACACCCAGGAGGCTGCTGGAATCTGACCTGCCCAA 120  
 DB 384 CCTGAGAGACATTCACACACCCAGGAGGCTGCTGGAATCTGACCTGCCCAA 325  
 QY 121 GAAATGTTGTCACAGCTGACAGAGATGACAGAGCTTAGTACGACATCAGAAATC 180  
 DB 324 GAAATGTTGTCACAGCTGACAGAGATGACAGAGCTTAGTACGACATCAGAAATC 265  
 QY 181 ATTCCTAATTTGTACAGAGCTGAAGAAACACCTGCTGAAGGGCAGCCATTTGAAG 240  
 DB 264 ATTCCTAATTTGTACAGAGCTGAAGAAACACCTGCTGAAGGGCAGCCATTTGAAG 205  
 QY 241 AAAATAATCTTCCCGCAAAAGAGAGTGGAGCTCAGAGTTGATTCATCTCTGTGT 300  
 DB 204 AAAATAATCTTCCCGCAAAAGAGAGTGGAGCTCAGAGTTGATTCATCTCTGTGT 145  
 QY 301 GAAGTAATTTGTGAGATGGAATCTCAGTTAAATATGTACA 342  
 DB 144 GAAGTAATTTGTGAGATGGAATCTCAGTTAAATATGTACA 103

RESULT 2  
 BX102762  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 ORGANISM  
 Homo sapiens  
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 1 (bases 1 to 621)  
 Ebert, L., Heil, O., Hennig, S., Neubert, P., Partsch, E., Peters, M., Radelof, U., Schneider, D. and Korn, B.  
 Human Unigeneset - RZPD3  
 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany

RZPD: IMAGp998E2118.  
 RZPDLIB; I.M.A.G.E. cDNA Clone Collection;  
 Human Unigeneset - RZPD3 (RZPDLIB NO.972)  
 http://www.rzpd.de/CloneCards/cgi-bin/showlib.pl.cgi/response?libNo=972 Contact: Ina Rolfs  
 RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH  
 Heubnerweg 6, D-14059 Berlin, Germany  
 Tel: +49 30 32639 101  
 Fax: +49 30 32639 111  
 www.rzpd.de

This clone is available royalty-free from RZPD;  
 contact RZPD (clone@rzpd.de) for further information. Seq primer:  
 M13r, Primer sequence: TTCACACAGAACACGATGAC.

# FEATURES

Location/Qualifiers  
 1..621  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGp998E2118 ; IMAGE:66428"  
 /sex="male"  
 /dev\_stage="20 week post conception fetus"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares fetal liver spleen INFLS"  
 /note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia) with a modified polylinker; Site 1: Pac I; Site 2: Eco RI; 1st strand cDNA was primed with a Pac I - oligo(dT) primer [5' AACTGGAGAAATTAATAAGATCTTTTTTTTTTTTTTTT 3'], double-stranded cDNA was ligated to Eco RI adaptors (Pharmacia), digested with Pac I and cloned into the Pac I and Eco RI sites of the modified pT73 vector. Library went through one round of normalization. Library constructed by Bento Soares and M.Fatima Bonaldo."  
 BASE COUNT 200 a 141 c 138 g 142 t  
 ORIGIN  
 Query Match 100.0%; Score 342; DB 13; Length 621;  
 Best Local Similarity 100.0%; Pred. No. 9.1e-93;  
 Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
 QY 1 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATGCCCATG 60  
 DB 170 GCAGAGCTGAGGGGATGTGTCGCCGATTTGGAAACACTTGTGTCATATGCCCATG 229  
 QY 61 CCTGAGAGACATTCACACACCCAGGAGGCTGCTGGAATCTGGAATCTGGACGTCCTCAA 120  
 DB 230 CCTGAGAGACATTCACACACCCAGGAGGCTGCTGGAATCTGGACGTCCTCAA 289  
 QY 121 GAAATGTTGTCACAGCTGACAGAGATGACAGAGCTTAGTACGACATCAGAAATC 180  
 DB 290 GAAATGTTGTCACAGCTGACAGAGATGACAGAGCTTAGTACGACATCAGAAATC 349  
 QY 181 ATTCCTAATTTGTGACGAGCTGAAGAAACACCTGCTGAAGGGCAGCCATTTGAAG 240  
 DB 350 ATTCCTAATTTGTGACGAGCTGAAGAAACACCTGCTGAAGGGCAGCCATTTGAAG 409  
 QY 241 AAAATAATCTTCCCGCAAAAGAGAGTGGAGCTCAGAGTTGATTCATCTCTGTGT 300  
 DB 410 AAAATAATCTTCCCGCAAAAGAGAGTGGAGCTCAGAGTTGATTCATCTCTGTGT 469  
 QY 301 GAAGTAATTTGTGAGATGGAATCTCAGTTAAATATGTACA 342  
 DB 470 GAAGTAATTTGTGAGATGGAATCTCAGTTAAATATGTACA 511

RESULT 3  
 AU136752  
 LOCUS  
 DEFINITION  
 ACCESSION  
 VERSION  
 KEYWORDS  
 SOURCE  
 Homo sapiens (human)  
 AU136752 816 bp mRNA linear EST 02-AUG-2002  
 AU136752 PLACE1 Homo sapiens cDNA clone PLACE1005004 5', mRNA sequence.  
 AU136752  
 AU136752.1 GI:10997291  
 EST.  
 Homo sapiens (human)

```

ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 816)
Ota,T., Nishikawa,T., Suzuki,Y., Ishii,S., Saito,K., Kawai,Y.,
Yamamoto,J., Wakamatsu,A., Nakamura,Y., Nagai,T., Sugano,S. and
Isogai,T.
TITLE HRI human cDNA project
JOURNAL Unpublished
COMMENT Contact: Takao Isogai
Genomics Laboratory
Helix Research Institute
1532-3 Yana, Kisarazu, Chiba 292-0812, Japan
Tel: 81-438-52-3975
Fax: 81-438-52-3986
Email: genomics@hri.co.jp
HRI human cDNA project; 5'- & 3'-end one pass sequencing; Helix
Research Institute; cDNA library construction; Department of
Virology, Institute of Medical Science, University of Tokyo, and
Helix Research Institute.
FEATURES
    source
        Location/Qualifiers
            1..816
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="PLACE1005004"
                /tissue_type="placenta"
                /clone_lib="PLACE1"
                /note="Vector: pME185FL3"
BASE COUNT 259 a 170 c 177 g 204 t 6 others
ORIGIN
    Query Match 100.0%; Score 342; DB 9; Length 816;
    Best Local Similarity 100.0%; Pred. No. 1e-92;
    Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCATG 60
Db 181 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCAIG 240

QY 61 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCCAAA 120
Db 241 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCCAAA 300

QY 121 GAATGGTGTCAACCTCCCAACAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 301 GAATGGTGTCAACCTCCCAACAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 360

QY 181 ATTCCTTAATTTGTACACAGAGCTGAAGAAACACITTCGTGAGGGCAGCCATCATTTGAAG 240
Db 361 ATTCCTTAATTTGTACACAGAGCTGAAGAAACACITTCGTGAGGGCAGCCATCATTTGAAG 420

QY 241 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCCTGTGT 300
Db 421 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCCTGTGT 480

QY 301 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 342
Db 481 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 522

RESULT 4
AL550042 1201 bp mRNA linear EST 31-MAY-2003
LOCUS AL550042 Homo sapiens PLACENTA COT 25-NORMALIZED Homo sapiens cDNA
DEFINITION clone CS0DI053YC21 5-PRIME, mRNA sequence.
ACCESSION AL550042
VERSION AL550042.2 GI:31271860
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

```

```

REFERENCE 1 (bases 1 to 1201)
AUTHORS Li,W.B., Gruber,C., Jessee,J. and Polayes,D.
TITLE Full-length cDNA libraries and normalization
JOURNAL Unpublished
COMMENT On Feb 15, 2001 this sequence version replaced gi:12886616.
Contact: Genoscope
Genoscope - Centre National de Sequencage
BP 191 91006 EVRY cedex - France
Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr
Library was constructed by Life Technologies, a division of
Invitrogen. This sequence belongs to sequence cluster 5302.f For
more information about this cluster, see
http://www.genoscope.cns.fr/cgi-bin/Cluster.cgi?seq=CS0DI053AB110P1&cluster=5302.f. Contact :
Feng Liang Email : fliang@lifetech.com URL :
http://fulllength.invitrogen.com/Invitrogen Corporation 1600
Faraday Avenue Genoscope sequence ID : CS0DI053AB110P1.
FEATURES
    source
        Location/Qualifiers
            1..1201
                /organism="Homo sapiens"
                /mol_type="mRNA"
                /db_xref="taxon:9606"
                /clone="CS0DI053YC21"
                /tissue_type="PLACENTA COT 25-NORMALIZED"
                /clone_lib="Homo sapiens PLACENTA COT 25-NORMALIZED"
                /note="1st strand cDNA was primed with a NotI-oligo(dT)
                primer. Five prime end enriched, double-strand cDNA was
                digested with Not I and cloned into the Not I and EcoR V
                sites of the pCMVSPORT 6 vector. Library was normalized."
BASE COUNT 373 a 243 c 240 g 321 t 24 others
ORIGIN
    Query Match 100.0%; Score 342; DB 9; Length 1201;
    Best Local Similarity 100.0%; Pred. No. 1.3e-92;
    Matches 342; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCATG 60
Db 223 GCAGAGCTGAGGGGATGTGGTCCCGGATTTGGAAACACITTCGTGTCATATGGCCCCATG 282

QY 61 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCCAAA 120
Db 283 CQTGAGAGACATTCACACACACCCAGGAGGTGGCTGGGAATCTGGAGCTGCCCAAA 342

QY 121 GAATGGTGTCAACCTCCCAACAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 180
Db 343 GAATGGTGTCAACCTCCCAACAACAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 402

QY 181 ATTCCTTAATTTGTACACAGAGCTGAAGAAACACITTCGTGAGGGCAGCCATCATTTGAAG 240
Db 403 ATTCCTTAATTTGTACACAGAGCTGAAGAAACACITTCGTGAGGGCAGCCATCATTTGAAG 462

QY 241 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCCTGTGT 300
Db 463 AAAATAATACTTTCCCGCAAAAAGAGAGTGGAGTCACAGATTTGATCCATTCCTGTGT 522

QY 301 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 342
Db 523 GAAGTAATTTGTGACGATGGAACTTCAGTTAAATATGTACA 564

RESULT 5
AL148843/c 478 bp mRNA linear EST 30-SEP-1998
LOCUS AL148843 qc65h02.x1 Soares.placenta.8to9weeks.2NDHP8to9W Homo sapiens cDNA
DEFINITION clone IMAGE:1714515 3' similar to gb:X59357 60S RIBOSOMAL PROTEIN
L22 (HUMAN);, mRNA sequence.
ACCESSION AL148843
VERSION AL148843.1 GI:3677312
KEYWORDS EST.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

```

Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 478)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 This clone is available royalty-free through LNL; contact the  
 IMAGE Consortium ([info@image.lnl.gov](mailto:info@image.lnl.gov)) for further information.  
 Seq primer: -40ml3 fwd. ET from Amersham  
 High quality sequence stop: 434.  
 Location/Qualifiers  
 1..478  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="IMAGE:1714515"  
 /dev\_stage="two placentae; one from 8 weeks and another  
 from 9 weeks post conception"  
 /lab\_host="DH10B (ampicillin resistant)"  
 /clone\_lib="Soares.placenta\_8to9weeks\_2NHP8to9W"  
 /note="Organ: Placenta; Vector: pT7T3D (Pharmacia) with a  
 modified polylinker; Site 1: Not I; Site 2: Eco RI; 1st  
 strand cDNA was primed with a Not I - oligo(dT) primer [5',  
 TGTACCAATCTGAAGTGGGAGCGCGGATTTTTTTTTTTT 3'],  
 double-stranded cDNA was size selected, ligated to Eco RI  
 adapters (Pharmacia), digested with Not I and cloned into  
 the Not I and Eco RI sites of a modified pT7T3 vector  
 (Pharmacia). Library constructed by Bento Soares and  
 M.Fatima Bonaldo."  
 BASE COUNT 119 a 103 c 105 g 151 t  
 ORIGIN  
 Query Match 99.5%; Score 340.4; DB 9; Length 478;  
 Best Local Similarity 99.7%; Pred. No. 2.4e-92;  
 Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCAGAGCTGAGGGATGTGTCCTCCCATTTGGAAACACTTGTCTATATGCCCCATG 60  
 DB 442 GCAGAGCTGAGGGATGTGTCCTCCCATTTGGAAACACTTGTCTATATGCCCCATG 383  
 QY 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGTCCCAAA 120  
 DB 382 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGTCCCAAA 323  
 QY 121 GAAATGCTGCACTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 180  
 DB 322 GAAATGCTGCACTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 263  
 QY 181 ATTCCTAATTTGTCACAGAGTGAAGAACCACTGCTCTGAAGGGCAGCCATCATTTGAAG 240  
 DB 262 ATTCCTAATTTGTCACAGAGTGAAGAACCACTGCTCTGAAGGGCAGCCATCATTTGAAG 203  
 QY 241 AATATATACTTTCCCGAAAAGAGAGTGGAGTCACAGATTTTCATCATCTGTTGT 300  
 DB 202 AATATATACTTTCCCGAAAAGAGAGTGGAGTCACAGATTTTCATCATCTGTTGT 143  
 QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342  
 DB 142 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 101  
 RESULT 6  
 BQ011259/c  
 LOCUS  
 DEFINITION  
 UI-1-BC0-afn-f-07-0-UI.s2 NCI-CGAP\_P11 Homo sapiens cDNA clone  
 UI-1-BC0-afn-f-07-0-UI 3', mRNA sequence.  
 ACCESSION  
 BQ011259  
 VERSION  
 BQ011259.1  
 KEYWORDS  
 EST.  
 SOURCE  
 Homo sapiens (human)  
 ORGANISM  
 Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 Mammalia; Euthera; Primates; Catarrhini; Hominiidae; Homo.  
 1 (bases 1 to 629)  
 NCI-CGAP <http://www.ncbi.nlm.nih.gov/ncicgap>.  
 National Cancer Institute, Cancer Genome Anatomy Project (CGAP),  
 Tumor Gene Index  
 Unpublished  
 Contact: Robert Strausberg, Ph.D.  
 Email: [cgaps-r@mail.nih.gov](mailto:cgaps-r@mail.nih.gov)  
 Tissue Procurement: Dr. Steven Brown  
 cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa  
 cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa  
 DNA Sequencing by: Dr. M. Bento Soares, University of Iowa  
 Clone Distribution: Clone distribution information can be obtained  
 from Dr. M. Bento Soares, [bento-soares@uiowa.edu](mailto:bento-soares@uiowa.edu)  
 Seq primer: M13 FORWARD  
 POLYA+yes.  
 Location/Qualifiers  
 1..629  
 /organism="Homo sapiens"  
 /mol\_type="mRNA"  
 /db\_xref="taxon:9606"  
 /clone="UI-1-BC0-afn-f-07-0-UI"  
 /tissue\_type="Placenta"  
 /dev\_stage="8-9 weeks"  
 /lab\_host="DH10B (Life Technologies)"  
 /clone\_lib="NCI-CGAP\_P11"  
 /note="Organ: Placenta; Vector: pT7T3-Pac (Pharmacia) with  
 a modified polylinker; Site 1: EcoR I; Site 2: Not I;  
 NCI-CGAP\_P11 is a cDNA library containing the following  
 tissue(s): Placenta 8-9 weeks pregnant. The library was  
 constructed according to Bonaldo, Lennon and Soares,  
 Genome Research, 6:791-806, 1996. First strand cDNA  
 synthesis was primed with an oligo-dT primer containing a  
 Not I site. Double stranded cDNA was ligated to an EcoR I  
 adaptor, digested with Not I, and cloned directionally  
 into pT7T3-Pac vector. The oligonucleotide used to prime  
 the synthesis of first-strand cDNA contains a library tag  
 sequence that is located between the Not I site and the  
 (dT)18 tail. The sequence tag for this library is GA.  
 TAG\_L18-UI-1-BC0  
 TAG\_TISSUE=placenta human 8 week  
 TAG\_SEQ=GA"  
 BASE COUNT 142 a 142 c 142 g 203 t  
 ORIGIN  
 Query Match 99.5%; Score 340.4; DB 12; Length 629;  
 Best Local Similarity 99.7%; Pred. No. 2.8e-92;  
 Matches 341; Conservative 0; Mismatches 1; Indels 0; Gaps 0;  
 QY 1 GCAGAGCTGAGGGATGTGTCCTCCCATTTGGAAACACTTGTCTATATGCCCCATG 60  
 DB 455 GCAGAGCTGAGGGATGTGTCCTCCCATTTGGAAACACTTGTCTATATGCCCCATG 396  
 QY 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGTCCCAAA 120  
 DB 395 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGAGTCCCAAA 336  
 QY 121 GAAATGCTGCACTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 180  
 DB 335 GAAATGCTGCACTCCACACCAAGATGGACAGCCTTAGTACGACATCAGAAATTC 276  
 QY 181 ATTCCTAATTTGTCACAGAGTGAAGAACCACTGCTCTGAAGGGCAGCCATCATTTGAAG 240  
 DB 275 ATTCCTAATTTGTCACAGAGTGAAGAACCACTGCTCTGAAGGGCAGCCATCATTTGAAG 216  
 QY 241 AATATATACTTTCCCGAAAAGAGAGTGGAGTCACAGATTTTCATCATCTGTTGT 300  
 DB 215 AATATATACTTTCCCGAAAAGAGAGTGGAGTCACAGATTTTCATCATCTGTTGT 156  
 QY 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 342  
 DB 155 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTAATGTACA 114

```

RESULT 7
AA777897/c
LOCUS
DEFINITION
2193605.s1 Soares fetal_liver_spleen.INFLS_S1 Homo sapiens cDNA
clone IMAGE:448352 3' similar to gb:X59357 60S RIBOSOMAL PROTEIN
L22 (HUMAN);, mRNA sequence.

ACCESSION
AA777897
VERSION
AA777897.1
KEYWORDS
GI:2836890
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
Hillier,D., Allen,M., Bowles,L., Dubuque,T., Gelsel,G., Jost,S.,
Krizman,D., Kucaba,T., Lacy,M., Le,N., Lennon,G., Marra,M., Martin
J., Moore,B., Schellenberg,K., Steptoe,M., Tan,F., Theising,B.,
White,Y., Wylie,T., Waterston,R. and Wilson,R.
WashU-NCI human EST Project
Unpublished
Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 515.
Location/Qualifiers
1..550
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="GDB:1352609"
/db_xref="taxon:9606"
/clone="IMAGE:448352"
/sex="male"
/dev_stage="20 week-post conception fetus"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_fetal_liver_spleen.INFLS_S1"
/note="Organ: Liver and Spleen; Vector: pT73D (Pharmacia)
with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
This is a subcloned version of the original Soares fetal
liver spleen INFLS library. 1st strand cDNA was primed
with a Pac I - oligo(dT) primer [5'
AAGTGAAGAATTAAATTAAGATCTTTTITTTTTTTTTTTT 3'],
double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Pac I and cloned into the Pac I
and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization. Library
constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT 136 a 123 c 127 g 164 t
ORIGIN
Query Match 96.8%; Score 331; DB 9; Length 550;
Best Local Similarity 99.7%; Pred. No. 1.9e-89;
Matches 342; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 GCAGAGCTGAGGGATGGTCCCGGATTTGGAAACACTTGTGTCATATTGCCCATG 60
|||||
DB 438 GCAGAGCTGAGGGATGGTCCCGGATTTGGAAACACTTGTGTCATATTGCCCATG 379
|||||

QY 61 CCTGAGAGACATTCACC - ACCACCCAGAGGGTGGCTGCTGGAATCTGAGCTCCCAA 119
|||||
DB 378 CCTGAGAGACATTCACC - ACCACCCAGAGGGTGGCTGCTGGAATCTGAGCTCCCAA 319
|||||

QY 120 AGAATGTGTCAACCTCCACACAAAGATGGACAGCCCTTAGGTACGACATCAGAAATT 179
|||||
DB 318 AGAATGTGTCAACCTCCACACAAAGATGGACAGCCCTTAGGTACGACATCAGAAATT 259
|||||

QY 180 CATTCCTAAATTTGCACAGAGCTGAAGAAACCACTGCTCTGAAGGGGAGCCATCATTGAA 239
|||||
DB 258 CATTCCTAAATTTGCACAGAGCTGAAGAAACCACTGCTCTGAAGGGGAGCCATCATTGAA 199
|||||

QY 240 GAAATATATCTTTCCCGCAAAAGAGAGTGGACGTACAGATTTTGCATCTCTCTTG 299
|||||
DB 198 GAAATATATCTTTCCCGCAAAAGAGAGTGGACGTACAGATTTTGCATCTCTCTTG 139
|||||

QY 300 TGAAGTAAATTTGTGACGATGGAACTTCAGTTAAATATATGACA 342
|||||
DB 138 TGAAGTAAATTTGTGACGATGGAACTTCAGTTAAATATATGACA 96
|||||

RESULT 8
AA128289/c
LOCUS
DEFINITION
qc56g12.x1 Soares_placenta_8to9weeks_2NbHP8to9w Homo sapiens cDNA
clone IMAGE:1714630 3' similar to gb:X59357 60S RIBOSOMAL PROTEIN
L22 (HUMAN);, mRNA sequence.

ACCESSION
AA128289
VERSION
AA128289.1
KEYWORDS
GI:3596803
SOURCE
Homo sapiens (human)
ORGANISM
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo..
1 (bases 1 to 448)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL ; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 655 Std Error: 0.00
Seq primer: -40ml3 fwd. ET from Amersham
High quality sequence stop: 443.
Location/Qualifiers
1..448
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:1714630"
/dev_stage="two placenta; one from 8 weeks and another
from 9 weeks post conception"
/lab_host="DH10B (ampicillin resistant)"
/clone_lib="Soares_placenta_8to9weeks_2NbHP8to9w"
/note="Organ: Placenta; Vector: pT73D (Pharmacia) with a
modified polylinker; Site_1: Not I; Site_2: Eco RI; 1st
strand cDNA was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGATTTTTTTTTTTT 3'],
double-stranded cDNA was size selected, ligated to Eco RI
adapters (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of a modified pT73 vector
(Pharmacia). Library constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 114 a 95 c 96 g 143 t
ORIGIN
Query Match 96.0%; Score 328.4; DB 9; Length 448;
Best Local Similarity 99.4%; Pred. No. 1e-88;
Matches 340; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1 GCAGAGCTGAGGGATGGTCCCGATTTGGAAACACTTGTGTCATATTGCCCATG 60
|||||
DB 441 GCAGAGCTGAGGGATGGTCCCGATTTGGAAACACTTGTGTCATATTGCCCATG 382
|||||

QY 61 CCTGAGAGACATTCACCACCACCCAGAGGGTGGCTGCTGGAATCTGAGCTCCCAA 120
|||||
DB 381 CCTGAGAGACATTCACCACCACCCAGAGGGTGGCTGCTGGAATCTGAGCTCCCAA 323
|||||

QY 121 GAATGGTGTCAACCTCCACAAAGATGGCAAGCCCTTAGGTACGACATCAGAAATTC 180
|||||

```

```

322 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATCAGAATTC 263
|||||
181 ATTCTAATTTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAITGAAG 240
|||||
262 ATTCTAATTTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAITGAAG 203
|||||
241 AAAATATATCTTTCCCGCAAAAGAGAAGTGGAGCTCAGACATTTGATGCCATTTCTGTGT 300
|||||
202 AAAATATATCTTTCCCGCAAAAGAGAAGTGGAGCTCAGACATTTGATGCCATTTCTGTGT 143
|||||
301 GAAGTAATTTGTGAGAGCGCACTTCACTTAATTAATATGACA 342
|||||
142 GAAGTAATTTGTGAGAGCGCACTTCACTTAATTAATATGACA 101
|||||

RESULT 9
AI800597/c
LOCUS      523 bp      mRNA      linear      EST 19-DEC-1999
DEFINITION w912405.x1 Soares NSF_F8_9W_OT_PA_P_S1 Homo sapiens cDNA clone
IMAGE:2364873 3' similar to SW:INL4_HUMAN Q14641 EARLY PLACENTA
INSULIN-LIKE PEPTIDE PRECURSOR ; mRNA sequence.
ACCESSION  AI800597
VERSION     AI800597
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 523)
NCI-CCGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
Insert Length: 1083 Std Error: 0.00
Seq primer: -40UP from GIBCO
High quality sequence stop: 451.
Location/Qualifiers
1. 523
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:2364873"
/lab_host="DH10B"
/clone_lib="Soares NSF_F8_9W_OT_PA_P_S1"
/note="Organ: pooled; Vector: pT73D-Pac (Pharmacia) with
a modified polylinker; Site_1: Not 1; Site_2: Eco RI;
Equal amounts of plasmid DNA from five normalized
libraries were mixed, and ss circles were made in vitro.
Following HAP purification, this DNA was used as tracer in
a subtractive hybridization reaction. The driver was
PCR-amplified cDNAs from pools of 5,000 clones made from
the same 5 libraries. The pools consisted of the following
libraries and clones: Soares NBHSF pool 1:
309384-310919, 323208-325895 Soares NB2HP pool 1:
145032-147335, 147720-148103, 148872-149255, 15002 -
150407, 151176-152327 Soares NB2HF8-9W pool 1:
758280-760583, 772104-774407 Soares NBHPA pool 1:
304776-306311, 320136-322823, 326280-326663 Soares NBHCT
pool 1: 723720-726407, 739080-740999 Subtraction by Bento
Soares and M. Fatima Bonaldo."
BASE COUNT 133 a 111 c 117 g 162 t
ORIGIN

Query Match 94.08; Score 321.4; DB 9; Length 523;
Best Local Similarity 98.08; Pred. No. 1.5e-86;
Matches 336; Conservative 0; Mismatches 6; Indels 1; Gaps 1;

1 GCAGAGCTGAGGGATGTGTCCTCCCGCATTTGGAAACACTTGTCTCATATTGCCCATG 60

```

```

. Db      |||||
448 GCAGAGCTGAGGGATGTGTCCTCCCGCATTTGGAAACACTTGTCTCATATTGCCCATG 389
|||||
61  CCTGGAACACATTTACACACACCCAGGAGGCTGCTGGAATCTGGAGCTGCCAAA 120
|||||
388 CCTGGAACACATTTACACACACCCAGGAGGCTGCTGGAATCTGGAGCTGCCAAA 329
|||||
121 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATC-AGAAAT 179
|||||
328 GAAATGGTGTCAACCTCCCAACAAAGATGGACAAGCCTTAGGTACGACATC-AGAAAT 269
|||||
180 CATTCTTAATTTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAATCAA 239
|||||
268 CATTCTTAATTTGTCCACGAGCTGAAGAAACCACTGTCTGAAGGCGACGCCATCAATCAA 209
|||||
240 GAAATATATCTTTCCCGCAAAAGAGAAGTGGAGCTCAGACATTTGATGCCATTTCTGTG 299
|||||
208 AAAATATATCTTTCCCGCAAAAGAGAAGTGGAGCTCAGACATTTGATGCCATTTCTGTG 149
|||||
300 TGAAGTAATTTGTGAGAGCGCACTTCACTTAATTAATATGACA 342
|||||
148 TGAAGTAATTTGTGAGAGCGCACTTCACTTAATTAATATGACA 106
|||||

RESULT 10
BG623890      948 bp      mRNA      linear      EST 18-APR-2001
LOCUS      602649042F1 NIH_MGC_79 Homo sapiens cDNA clone IMAGE:4770395 5',
DEFINITION mRNA sequence.
ACCESSION  BG623890
VERSION     BG623890.1 GI:13675261
KEYWORDS    EST.
SOURCE      Homo sapiens (human)
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 948)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LUCM1637 row: d column: 12
High quality sequence stop: 624.
Location/Qualifiers
1. 948
/organism="Homo sapiens"
/mol_type="mRNA"
/db_xref="taxon:9606"
/clone="IMAGE:4770395"
/lab_host="NIH_MGC_79"
/note="Organ: placenta; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccattatggcc); Site_2: SfiI (ggccattatggcc)
); 5' and 3' adaptors were used in cloning as follows: 5'
adaptor sequence: 5'-CACGGCCATTTATGCC-3' and 3' adaptor
sequence: 5'-ATTCTAGAGCGGAGCGCGGACATG-dt(30)BN-3'
(where B = A, C, G, or T). Average
insert size 1.3 kb (range 0.5-4.0 kb). 15/15 colonies
contained inserts by PCR. This library was enriched for
full-length clones and was constructed by Clontech
Laboratories (Palo Alto, CA). Note: this is a NIH_MGC
Library."
BASE COUNT 329 a 201 c 254 g 164 t
ORIGIN

```

```

Query Match          92.3%; Score 315.8; DB 10; Length 948;
Best Local Similarity 98.8%; Pred. No. 1e-84;
Matches 339; Conservative 0; Mismatches 2; Indels 2; Gaps 2;

Qy 1 GCAGAGCTGAGGGGATGCTGGTCCCGGATTTGGAAAACTTGTCTGCATATTTGCCCCCATG 60
Db 146 GCAGAGTTGAAGGGATGCTGGTCCCGGATTTGSAANAAGGATGCTGTCTATATTTGCCCCCATG 205

Qy 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACGTCGCCAAA 120
Db 200 GCAGAGCTGAGGGGATGCTGGTCCCGGATTTGGAAAACTTGTCTGCATATTTGCCCCCATG 259

Qy 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACGTCGCCAAA 120
Db 260 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACGTCGCCAAA 318

Qy 121 GAA-ATGGTGTCAACCTCCCAACAAGATGACAAAGCTTAGGTACGACATCAGAATT 179
Db 319 GACATGGTGTCAAGCTCCCAACAAGATGACAAAGCTTAGGTACGACATCAGAATT 378

Qy 180 CATTCTTAATTTGTCACCAAGCTGAAGAACCACTGTCTGAAGGGCAGCCCATCATTTGAA 239
Db 379 CATTCTTAATTTGTCACCAAGCTGAAGAACCACTGTCTGAAGGGCAGCCCATCATTTGAA 438

Qy 240 GAAATAAATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTTCTGTTG 299
Db 439 GAAATAAATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTTCTGTTG 498

Qy 300 TGAAGTAAATTTGACGATGAAGAACTTCAGTTAAATTTATGTACA 342
Db 499 TGAAGTAAATTTGACGATGAAGAACTTCAGTTAAATTTATGTACA 541

RESULT 11
CB549181
LOCUS          684 bp      mRNA      linear      EST 01-JUN-2003
DEFINITION    MPMPL0003_F01 MPMPL Macaca mulatta cDNA, mRNA sequence.
ACCESSION     CB549181
VERSION       CB549181.1 GI:31298376
KEYWORDS      EST.
SOURCE        Macaca mulatta (rhesus monkey)
ORGANISM      Macaca mulatta
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
              Cercopithecinae; Macaca.
REFERENCE     1 (bases 1 to 684)
AUTHORS       Katze,M.G., Bungarner,R., Korth,M., Feldman,R., Amjadi,M. and
              Holzman,T.
TITLE         Expressed sequence tags from Rhesus macaque placenta
JOURNAL       Unpublished
COMMENT       Contact: Holzman T
              Katze Lab
              University of Washington
              Box 358070, Seattle, WA 98195-8070, USA
              Tel: 206 732 6156
              Fax: 206 732 6055
              Email: ted@locke.hs.washington.edu
              Similar to GenBank entry BC026254 Homo sapiens,
              insulin-like 4 (placenta), clone MGC:22499 IMAGE:4770395, mRNA,
              complete cds. 4/2002
              Plate: MPMPL0003 row: F column: 01.
              Location/Qualifiers
                1..684
                  /organism="Macaca mulatta"
                  /mol_type="mRNA"
                  /db_xref="taxon:9544"
                  /sex="male"
                  /dev_stage="neonatal"
                  /clone_lib="MMPL"
                  /note="Organ: placenta"
BASE COUNT    215 a 147 c 150 g 172 t
ORIGIN

```

```

Qy 1 GCAGAGCTGAGGGGATGCTGGTCCCGGATTTGGAAAACTTGTCTGCATATTTGCCCCCATG 60
Db 146 GCAGAGTTGAAGGGATGCTGGTCCCGGATTTGSAANAAGGATGCTGTCTATATTTGCCCCCATG 205

Qy 61 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACGTCGCCAAA 120
Db 206 CCTGAGAAGACATTCACCAACCCAGAGGGTGGCTGCTGGAATCTGGACGTCGCCAAA 265

Qy 121 GAAATGTTGTCAACCTCCCAACAAGATGACAAAGCTTAGGTACGACATCAGAATTC 180
Db 266 GAAACGGTGTCAACCTCCCAACAAGATGACAAAGCTTAGGTACGACATCAGAATTC 325

Qy 181 ATTCTTAATTTGTCACCAAGCTGAAGAACCACTGTCTGAAGGGCAGCCCATCATTTGAG 240
Db 326 ATTCTTAATTTGTCACCAAGCTGAAGAACCACTGTCTGAAGGGCAGCCCATCATTTGAG 385

Qy 241 AAAATAATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTTCTGTTG 300
Db 386 AAAATAATCTTTCCCGCAAAAAGAGAGTGGACGTCACAGATTTGATCCATTTCTGTTG 445

Qy 301 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 342
Db 446 GAAGTAATTTGTGACGATGGAACCTTCAGTTAAATTTATGTACA 487

RESULT 12
CB548981
LOCUS          701 bp      mRNA      linear      EST 01-JUN-2003
DEFINITION    MPMPL0017_E12 MPMPL Macaca mulatta cDNA, mRNA sequence.
ACCESSION     CB548981
VERSION       CB548981.1 GI:31298176
KEYWORDS      EST.
SOURCE        Macaca mulatta (rhesus monkey)
ORGANISM      Macaca mulatta
              Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
              Mammalia; Eutheria; Primates; Catarrhini; Cercopithecoidea;
              Cercopithecinae; Macaca.
REFERENCE     1 (bases 1 to 701)
AUTHORS       Katze,M.G., Bungarner,R., Korth,M., Feldman,R., Amjadi,M. and
              Holzman,T.
TITLE         Expressed sequence tags from Rhesus macaque placenta
JOURNAL       Unpublished
COMMENT       Contact: Holzman T
              Katze Lab
              University of Washington
              Box 358070, Seattle, WA 98195-8070, USA
              Tel: 206 732 6156
              Fax: 206 732 6055
              Email: ted@locke.hs.washington.edu
              Similar to GenBank entry BC026254 Homo sapiens,
              insulin-like 4 (placenta), clone MGC:22499 IMAGE:4770395, mRNA,
              complete cds. 4/2002
              Plate: MPMPL0017 row: E column: 12.
              Location/Qualifiers
                1..701
                  /organism="Macaca mulatta"
                  /mol_type="mRNA"
                  /db_xref="taxon:9544"
                  /sex="male"
                  /dev_stage="neonatal"
                  /clone_lib="MMPL"
                  /note="Organ: placenta"
BASE COUNT    221 a 145 c 156 g 179 t
ORIGIN

```

```

Query Match          90.2%; Score 308.4; DB 14; Length 701;
Best Local Similarity 93.9%; Pred. No. 1.6e-82;
Matches 321; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

Qy 1 GCAGAGCTGAGGGGATGCTGGTCCCGGATTTGGAAAACTTGTCTGCATATTTGCCCCCATG 60
Db 146 GCAGAGTTGAAGGGATGCTGGTCCCGGATTTGSAANAAGGATGCTGTCTATATTTGCCCCCATG 205

```